An axiomatic approach to the ordinal number system

Ineke van der Berg

Thesis presented in partial fulfilment of the requirements for the degree of

Master of Science (Mathematics)

in the Faculty of Science at Stellenbosch University

Supervisor: Professor Z. Janelidze

2021

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: 6th March 2021

Acknowledgements

I would like to thank my supervisor, Professor Zurab Janelidze, not only for his guidance and contributions to the thesis, but also for nurturing my confidence and enthusiasm throughout my studies. I could not imagine a better guide through the magical realm of mathematics.

I also wish to thank my friend, Dr Katrina du Toit, for her skilful proofreading, usually on short notice. I value your appreciation for brevity and precision, and your feedback is exactly the right level of pedantic. I promise to use more commas in the future.

Thanks, also, to my other friends, particularly Roy, Alan, and Dale, for their comments on topology, typography, and layout, respectively.

Iggy, thank you for being supportive, for motivating me, and for always being available for conversation, whether academic or otherwise.

Finally, to my parents, thank you for your constant encouragement and support, for putting up with me during the lockdown, and for feeding me whenever I needed it.

Abstract

An axiomatic approach to the ordinal number system

I. van der Berg

Department of Mathematical Sciences,
University of Stellenbosch,
Private Bag X1, 7602 Matieland, South Africa.

Thesis: MSc (Mathematics)

2021

Ordinal numbers are transfinite generalisations of natural numbers, and are usually defined and studied concretely as special types of sets. In this thesis we explore an abstract approach to developing the theory of ordinal numbers, where we present various axiomatisations of an ordinal number system and prove their equivalence. Since ordinal numbers do not form a set, in order to develop such a theory one needs to extend the usual framework of Zermelo-Fraenkel set theory. Among several such possible extensions, we pick the one that is based on the notion of a Grothendieck universe. While some of the results obtained in this thesis are merely adaptations of known results to this context, some others are new even to classical set theory. Among these is a definition and a universal property of the ordinal number system that mimics the classical Dedekind-Peano approach to the natural number system.

Opsomming

’n Aksiomatiese benadering tot die ordinaalgetalstelsel

(“An axiomatic approach to the ordinal number system”)

I. van der Berg

Departement Wiskundige Wetenskappe,
Universiteit van Stellenbosch,
Privaatsak X1, 7602 Matieland, Suid Afrika.

Tesis: MSc (Wiskunde)

2021

Ordinaalgetalle is transfiniete veralgemenings van die telgetalle, en word gewoonlik konkreet gedefinieer en bestudeer as spesiale soorte stelle. In hierdie tesis ondersoek ons ’n abstrakte benadering tot die ontwikkeling van ordinaalgetalteorie, waarin ons verskeie aksiomatiserings van ordinaalgetalstelsels gee, en hul ekwivalensie bewys. Aangesien ordinaalgetalle nie ’n stel vorm nie, is dit nodig om die standaard raamwerk van Zermelo-Fraenkel stelteorie uit te brei om so ’n teorie te kan ontwikkel. Vanuit verskeie moontlike raamwerke kies ons een wat op die idee van ’n Grothendieck universum gebaseer is. Alhoewel sommige van die bevindings in hierdie tesis slegs aanpassings van bekende bevindings na hierdie konteks is, is ander nuut selfs in klassieke stelteorie. Die nuwe bevindings sluit ’n definisie en universele eienskap van die ordinaalgetalstelsel in, wat die klassieke Dedekind-Peano benadering tot die telgetalstelsel naboots.

Contents

Introduction

Set theory is a fascinating field of mathematics that has deep philosophical significance, as well as practical applications in almost every branch of mathematics. What makes set theory a 'theory', rather than merely a 'language', is essentially the theory of ordinal numbers and cardinal numbers, which are transfinite generalisations of the natural numbers. The founding steps in the study of ordinal and cardinal numbers were made by Georg Cantor, who is considered the father of this field of mathematics. It should be mentioned, however, that he was not alone in the endeavor of building this 'paradise' of set theory.[1] For instance, he shared many of his ideas with Richard Dedekind, who made significant contributions to set-theoretic thinking in mathematics.

The natural numbers are the 'counting numbers': $0, 1, 2, 3$, etc. One of the uses of natural numbers in mathematics – and in everyday life – is to measure the size of a set. Since natural numbers are finite, such a set must necessarily be finite. If we want to measure the size of an infinite set, we must surpass, in our imagination, all of the natural numbers. This is when the 'ordinal numbers' follow. It is an oddity of the transfinite world that many ordinal numbers measure the same size, which gives rise to a separate notion of a 'cardinal number' that gives a unique measure of an infinite set. Cardinal numbers can be defined independently of ordinal numbers (which was Cantor's original approach), or as particular types of ordinal numbers. The second approach turns out to be more powerful and is adopted in contemporary texts of set theory. With this approach, the theory of cardinal numbers becomes fundamentally intertwined with the theory of ordinal numbers.

One of the practical applications of set theory in mathematics is that it equips us with a formal language for generalising concrete mathematical objects to abstract mathematical structures. Interestingly, such a generalised approach to the system of ordinal numbers has not received much attention in the literature. In this thesis, we attempt to fill some of the gaps in the theory of ordinal numbers that were left by this

[1] David Hilbert wrote of set theory, "From the paradise which Cantor created for us, no-one shall be able to expel us." [1]

lack of attention.

To explain the nature of these gaps, let us first go back to natural numbers. In set theory, natural numbers can be defined concretely as particular types of ordinal numbers. However, Dedekind gave an alternative approach already in 1888 [2], where natural numbers are defined as objects in a 'natural number system' – an abstract mathematical structure fulfilling certain axioms, which are known today as Peano axioms (see [3]).

Since ordinal numbers extend the natural numbers, a natural question arises: what would a corresponding abstract approach to the ordinal system be? The ultimate goal of this thesis is to answer this question. However, before we can even formulate the question mathematically, some preliminary work is necessary. It turns out that the usual axiomatic set theory, as initiated by Ernst Zermelo in [4], might not be equipped for dealing with a question of this nature. The point of an axiomatic approach to set theory, in general, is to limit the domain of discussion to 'allowable' sets – the assumption that any collection that can be described by a formula forms a set leads to a contradiction, as is famously demonstrated by 'Russell's paradox'. The problem is that ordinal numbers do not form a set in the standard treatment of the subject – no matter how vast a set of ordinals is, one can always find an ordinal that is bigger than all of them. In standard set theory this is usually not a problem, since the focus of study there is not on a single object representing the totality of ordinal numbers, but on individual ordinal numbers, which are indeed sets. In what we aim to achieve, however, not only do we want to view all ordinal numbers as elements of some particular set, but we even want to consider and compare many such sets with one another. Among the results that we obtain in this thesis are two 'universal properties' of the ordinal number system, each of which presents an ordinal number system as an initial object in a suitable category. If we had been working in the standard axiomatic set theory, these objects would not necessarily have been sets. This means that this category itself has a size that is one level bigger than what is allowed, and thus, the standard trick of set theory to talk about 'classes' when we need to work with larger collections of sets would fail, since it is not possible to consider a 'class of classes'.

The size issues mentioned in the previous paragraph can be fixed by a method similar to the one used by Saunders Mac Lane in his exposition of category theory [5]. There are alternative methods, such as the axiomatic Gödel-Bernays set theory, but since the research in this thesis was largely motivated by a search for category-theoretic universal properties of the ordinal number system, we chose to follow Mac Lane's approach. In this approach, we develop the theory of ordinal numbers within, or rather, relative to a miniature version of the context of set theory, called a 'universe'. The idea comes from Alexander Grothendieck [6], although we do not adopt his 'axiom of universes' and we slightly relax the notion of a universe as he introduced it.

This thesis is organised in four chapters. Chapter 1 recalls the basic material from set theory needed to formulate and prove the results contained in the rest of the thesis. In the same chapter we develop the necessary preliminaries on universes.

Chapter 2 develops an abstract theory of ordinal numbers, leading to a universal property. Although the proof of this universal property is easy, the task of finding its formulation was non-trivial (at least it was the case for my supervisor, who did not succeed and was very happy with the outcome of my attempt). It turned out, however, that this was a rediscovery of a known universal property. In the work of André Joyal and Ieke Moerdijk on algebraic set theory [7], the same universal property is used as a possible definition of the ordinal number system, although in a slightly different axiomatic setup from ours.

In Chapter 3, we reach the ultimate goal of the thesis as mentioned above: we give an entirely new account of abstract ordinal number systems, which closely follows Dedekind's approach to abstract natural number systems. This leads to a different universal property, which, unlike the one in Chapter 2, is not found in algebraic set theory. The principal difference between these approaches is that the latter does not rely on the successive order of ordinal numbers as the base structure (i.e. on 'well-ordering'), but rather derives the order as a consequence of a Dedekind-style structure. Interestingly, this derivation begins by defining an Alexandrov topology. The proof of the equivalence of the axiomatisation proposed in Chapter 3 and the more standard notion of an ordinal system as defined in Chapter 2 is, in fact, the most involved proof of the thesis.

Finally, in Chapter 4, we make some remarks on possible future research on the theme of this thesis, and we conclude with a topic that has served as the initial inspiration for the work contained in this thesis.

Slightly more detailed descriptions of the chapters are given in their first few paragraphs.

Some of the results contained in this thesis have been written up in a paper [8], coauthored with my supervisor, which is at the moment submitted for publication. In particular, the paper is based on Section 1.5, Sections 2.1–2.4, and Chapter 3. The material presented in these sections is generally an expanded version of the material found in the paper, but in a number of instances the content of the thesis is identical to that of the paper.

I have given the following talks on some of the results of the thesis:

- 'An axiomatic approach to ordinal numbers', at the Annual Congress of the South African Mathematical Society, University of Cape Town, 4 December 2019.
- 'Incremented joins and the ordinal number system', at the UJ Logic Workshop, University of Johannesburg, 28 January 2020.

The associative addition of sets in Section 4.3 is actually based on an idea I had in the second year module 'Foundations of Abstract Mathematics I', in which I was introduced to set theory. We were asked to come up with a way of 'adding' sets such that the cardinality of the sum of two sets equals the sum of their cardinalities. I proposed the following: take the elements of the first term, and recursively insert them into each node in the 'tree' of elements of the second term (i.e. into each element of each element of each element of the second term, etc.) – see Figure 1 and Figure 2.

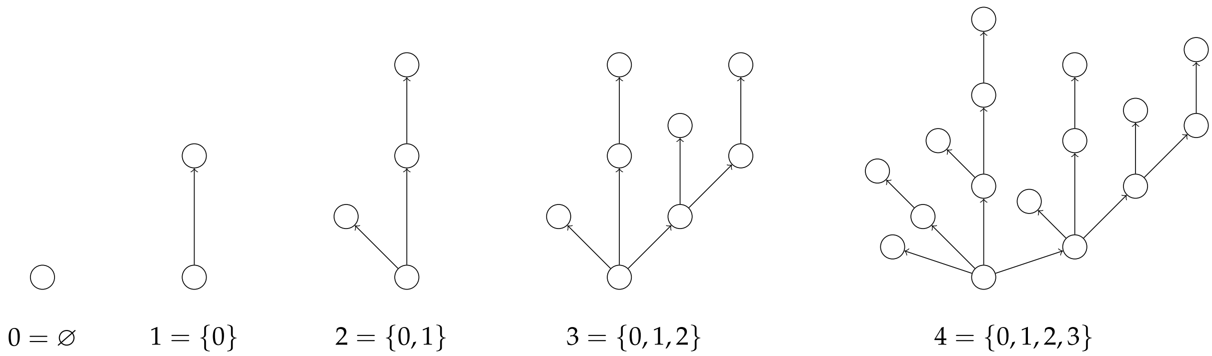


Figure 1: The first four von Neumann ordinals, seen as rooted trees.

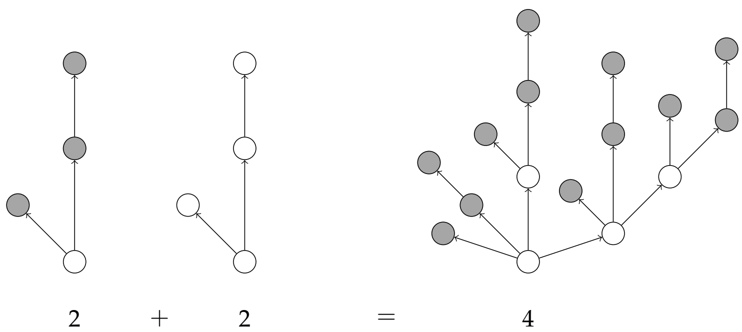


Figure 2: Addition performed by recursively inserting each element of the first term into each node in the graph of the second term. This figure illustrates that $2 + 2 = 4$ when the terms are von Neumann natural numbers.

This was not the solution the lecturer – now my supervisor – expected. In fact, he was not familiar with it at all. While for ordinal numbers my addition matched with the usual ordinal addition, for general sets the construction seemed to be new. What is remarkable about this way of adding sets is its property of being strictly associative, unlike the sum of sets given by the usual disjoint union, which is only associative up to bijection. In order to formalise this construction beyond finite sets, I first needed to study the theory of ordinal numbers. Soon after embarking on such a study and even before commencing with my MSc research, we discovered that the addition of sets I had found was presented by Alfred Tarski at a conference in 1955 [9], although he never published this work. We learnt this from a paper [10] that aims to recover the unpublished work of Tarski and explore this idea further.

Chapter 1

Background on set theory

This chapter is devoted to the set-theoretic background that is necessary for the development of the thesis.

The language of Zermelo-Fraenkel set theory is outlined in Section 1.1. The discussion of formulas in that section is loosely based on [11], while the brief overview of models is based on the definitions in [12].

Section 1.2 gives the Zermelo-Fraenkel axioms for set theory. These axioms were originally formulated by Ernst Zermelo in 1908 [4], with notable additions by Abraham Fraenkel in 1922 [13] and by John von Neumann in 1925 [14]. However, this thesis follows the notational conventions for sets and classes used in [15].

We will make extensive use of ordered sets, and thus some relevant definitions are given in Section 1.3, based on those in [16].

Since ordinal number systems extend natural number systems, the abstract natural number system of Richard Dedekind [2] and the concrete natural number system of von Neumann [14] will be discussed briefly in Section 1.4.

In Section 1.5 we investigate 'universes', relative to which we will develop our theory of ordinal numbers in subsequent chapters. Our universes are based on those introduced by Alexander Grothendieck in [6], although we relax one of his axioms.

Since the content of the first three sections of this chapter is fairly standard material, which can be found in almost every textbook of set theory, we do not include the proofs of any of the results mentioned there. The last section is complete with proofs, but, except for the presentation, we do not claim any originality of the results contained therein.

1.1 Language

This section specifies the notation that we will use in the formulas of the first-order language of set theory. These formulas are defined recursively as described below, using the notation in Table 1.1. We assume that the reader is familiar with first-order languages in general.

Terms: The terms of the language of set theory are all variables, one for each natural number: v_0, v_1, v_2, etc. They will be used to refer to sets. For notational convenience, we will often use other symbols, such as x, y, X, Y, etc., in the place of $v_0, v_1, v_2, \ldots$.

Predicates: The 'predicate symbols' used to define relations are $=$ (for equality), and $\in$ (for set membership).

Atomic formulas: If v_i and v_j are terms, then $v_i = v_j$ and $v_i \in v_j$ are formulas.

Connectives: If φ and ψ are formulas, then the following are also formulas:

$$\varphi \wedge \psi, \qquad \varphi \vee \psi, \qquad \neg\varphi, \qquad \varphi \Rightarrow \psi, \qquad \varphi \Leftrightarrow \psi.$$

Quantifiers: If v_i is a variable and φ is a formula, then $\forall_{v_i}\varphi$ and $\exists_{v_i}\varphi$ are formulas.

Table 1.1: The ingredients of a formula

Type	Name	Symbol	Natural language
Terms	Variables	$v_0, v_1, v_2, \ldots$	
Predicates	Equality	$=$	'equals'
	Membership	$\in$	'is an element of'
Connectives	Conjunction	$\wedge$	'and'
	Disjunction	$\vee$	'or'
	Negation	$\neg$	'not'
	Implication	$\Rightarrow$	'implies'
	Equivalence	$\Leftrightarrow$	'if and only if'
Quantifiers	Universal quantifier	$\forall$	'for all'
	Existential quantifier	$\exists$	'there exists'

Recall that a formula without free variables is a 'sentence'. If a formula φ has free variables, we write $\varphi(v_1, \ldots, v_n)$ to indicate that the free variables of φ are among $v_1, \ldots, v_n$.

Note that all terms in the language of set theory are variables – there are no function symbols or constant terms in this language. Instead, the functions and constants of set theory are special kinds of sets, which we will describe later.

Semantically, $\wedge$ and $\neg$ (or $\vee$ and $\neg$) can be used to construct all of the other connectives, and, given the connectives, $\forall$ can be used to construct $\exists$ (and vice versa).

If we are given a formula $\varphi(x, v_1, \ldots, v_n)$, along with terms $v_1, \ldots, v_n$, then, intuitively, we should be able to select precisely those sets x such that $\varphi(x, v_1, \ldots, v_n)$ holds for the specified values of $v_1, \ldots, v_n$. We use the notation

$$\{x \mid \varphi(x, v_1, \ldots, v_n)\}$$

to signify such a selection. We call the terms $v_1, \ldots, v_n$ that we have specified *parameters*.

This raises a question, however: can we form a set whose elements are exactly those sets selected by an arbitrary formula $\varphi(x, v_1, \ldots, v_n)$ with given parameters $v_1, \ldots, v_n$? The answer is no, in general, as famously illustrated by Russell's paradox: there can be no set that has as its elements exactly those sets x that satisfy $\neg[x \in x]$ (see e.g. [17]). If we try to imagine such a set $X = \{x \mid \neg[x \in x]\}$, we find that $X \in X$ if and only if $\neg[X \in X]$, which is obviously not permissible.

Russell's paradox motivates the idea of a 'class' – since the contradiction arises from the assumption that X is a set, it can be resolved by allowing X to be something other than a set. We achieve this by extending the $\in$ and $=$ notation so that sets can be seen as 'elements' of classes and classes can be seen as 'equal'. Thus, if a set x satisfies a formula $\varphi(x, v_1, \ldots, v_n)$ for given parameters $v_1, \ldots, v_n$, we write

$$x \in \{x \mid \varphi(x, v_1, \ldots, v_n)\},$$

and call $\{x \mid \varphi(x, v_1, \ldots, v_n)\}$ the *(definable) class* of all sets x that satisfy $\varphi(x, v_1, \ldots, v_n)$ for these parameters. If, for given parameters $v_1, \ldots, v_n$ and formulas $\varphi(x, v_1, \ldots, v_n)$ and $\psi(x, v_1, \ldots, v_n)$, it holds that $\varphi(x, v_1, \ldots, v_n) \Leftrightarrow \psi(x, v_1, \ldots, v_n)$, we write

$$\{x \mid \varphi(x, v_1, \ldots, v_n)\} = \{x \mid \psi(x, v_1, \ldots, v_n)\}.$$

While classes provide a useful way to talk about selections of sets, it is important to point out that we are perpetrating abuse of notation. If X is a proper class, it is never an element of itself, but neither does it satisfy $\neg[X \in X]$. The reason for this is that variables in the language of set theory are placeholders for sets, and thus the first-order predicate symbols $\in$ and $=$ cannot technically apply to anything other than sets. This is exactly what frees $X = \{x \mid \neg[x \in x]\}$ of Russell's paradox when X is a class, but it is, strictly speaking, more accurate to think of X simply as the formula $\neg[x \in x]$, and of the statement $x \in X$ simply as the claim that $\neg[x \in x]$ holds for the specified set x.

Since $x = x$ holds for all sets x, we can express the *class of all sets* as follows:

$$V = \{x \mid x = x\}.$$

If, for given parameters $v_1, \ldots, v_n$, it holds that

$$\psi(x, v_1, \ldots, v_n) \quad \Rightarrow \quad \varphi(x, v_1, \ldots, v_n),$$

we call the class $\{x \mid \psi(x, v_1, \ldots, v_n)\}$ a *subclass* of $\{x \mid \varphi(x, v_1, \ldots, v_n)\}$.

If, for given parameters $v_1, \ldots, v_n$, a set y exists such that

$$x \in y \quad \Leftrightarrow \quad x \in \{x \mid \varphi(x, v_1, \ldots, v_n)\},$$

we write

$$y = \{x \mid \varphi(x, v_1, \ldots, v_n)\},$$

and say that the formula $\varphi(x, v_1, \ldots, v_n)$ *defines* a set and the class $\{x \mid \varphi(x, v_1, \ldots, v_n)\}$ *forms* a set for the given parameters. A class that does not form a set is called a *proper* class.

If a class X is a subclass of a class Y, and X and Y each forms a set, we write $X \subseteq Y$ and call X a *subset* of Y – equivalently,

$$X \subseteq Y \quad \Leftrightarrow \quad \forall_x [x \in X \Rightarrow x \in Y].$$

If $X \subseteq Y$ and $X \neq Y$, we write $X \subset Y$ and call X a *proper* subset of Y.

A binary *class relation* R for given parameters $v_1, \ldots, v_n$ is a class

$$\begin{aligned} R &= \{z \mid \exists_{x,y}[z = (x, y) \wedge \varphi(z, v_1, \ldots v_n)]\} \\ &= \{(x, y) \mid \varphi(x, y, v_1, \ldots v_n)\}. \end{aligned}$$

We write $x \; R \; y$ to indicate $(x, y) \in R$, and $x \not R \; y$ to indicate $\neg[(x, y) \in R]$ (we will explain the 'ordered pair' notation (x, y) in Section 1.2). A binary class relation that forms a set is simply called a binary *relation*.

A *class function*, for given parameters $v_1, \ldots, v_n$, is a binary class relation F such that for all sets x, y_1, y_2,

$$[(x, y_1) \in F \wedge (x, y_2) \in F] \quad \Rightarrow \quad [y_1 = y_2].$$

A class function f that forms a set is simply called a *function*. If F is a class function (or relation), the *domain* and *range* of F are, respectively,

$$\operatorname{dom} F = \{x \mid \exists_y[(x,y) \in F]\},$$
$$\operatorname{range} F = \{y \mid \exists_x[(x,y) \in F]\}.$$

For a class function F, if $(x,y) \in F$, we write $F(x) = y$ and call y the *value* of F at x.

If the domain of a function f is $\operatorname{dom} f = X_1 \times \ldots \times X_n$ (the 'product' notation $\times$ is explained in Section 1.2), we call f an *n-ary* function and write $f(x_1, \ldots, x_n)$ for the value of f at $(x_1, \ldots, x_n)$.

The *image* of a class X under a class function F is the class

$$FX = \{F(x) \mid x \in X \wedge x \in \operatorname{dom} F\}.$$

More generally, the image of a class X under a binary class relation R is the class $\{y \mid \exists_{x \in X}\ x\ R\ y\}$.

When referring to a class function, we may write $F\colon \operatorname{dom} F \to Y$, where $\operatorname{range} F$ is a subclass of Y, and say that F is a class function from $\operatorname{dom} F$ to Y, and that Y is the *codomain* or *set of destination* of F. If $\operatorname{dom} F$ is a subclass of a class X, we write $F\colon X \to Y$ and say that F is a *partial* class function from X to Y.

We say a class function $F\colon \operatorname{dom} F \to Y$ is *surjective* and call it a *surjection* if $Y = \operatorname{range} F$. We say it is *injective* and call it an *injection* if, for all sets x_1, x_2, y,

$$[(x_1,y) \in F \wedge (x_2,y) \in F] \quad \Rightarrow \quad [x_1 = x_2].$$

If F is both surjective and injective, we say it is *bijective* and call it a *bijection*.

We make a few exceptions to the convention of writing $F(x)$ for the value of F at x and FX for the image of X under F. In some cases parentheses in the image notation are necessary to avoid ambiguity, for instance, when a distinction must be made between $f(X \cup Y)$ and $fX \cup Y$. In general, we reserve the right to add parentheses whenever they add clarity.

Conversely, we omit parentheses for certain class functions when it improves visual clarity without creating ambiguity. We list them here.

(a) The 'power set' class function P, which maps a set X to the set $\mathsf{P}X$ consisting of all subsets of X.

(b) The 'restricted power set' function $\mathsf{P}_{\mathfrak{U}}$, explained in Section 1.5.

(c) The 'union' $\bigcup X$ of a class X of sets, as explained in Section 1.2, and its counterpart, the 'intersection' $\bigcap X$ of X.

(d) The following partial functions $\mathsf{P}X \to X$, for a partially ordered set X:

(i) $\mathsf{min}\, S$ and $\mathsf{max}\, S$, which give the smallest and largest element of a subset S of X, respectively;

(ii) $\bigvee S$ and $\bigwedge S$, which give the 'join' and 'meet' of a subset S of X, respectively, as explained in Section 1.3;

(iii) $\overset{+}{\bigvee} S$, which gives the 'incremented join' of a subset S of X, as explained in Section 2.2. Other 'collecting functions' $\overset{\circ}{\bigvee}$ also use this convention.

(e) The rank function, explained in Section 4.3.

We use the following abbreviations:

(a) For a subclass of a class X:

$$\{x \in X \mid \varphi(x, v_1, \ldots v_n)\} = \{x \mid [x \in X] \wedge \varphi(x, v_1, \ldots v_n)\}.$$

(b) For a set with a finite number of elements:

$$\{x_1, x_2, \ldots, x_n\} = \{y \mid [y = x_1] \vee [y = x_2] \vee \ldots \vee [y = x_n]\}.$$

Finally, we need to discuss the concept of a 'model'. Since it is not the subject of this thesis, we do not set out to make the definition of a model precise here. Instead, we will try to provide an intuition for the concept, which will allow us to make some interesting remarks.

Let M be a class and let φ be a formula. Then the *relativisation* of φ to M (written φ^M) can be built up recursively from φ, as in Table 1.2.

Table 1.2: Relativisation of a formula

Notation	Interpretation
$[x = y]^M$	$x = y$
$[x \in y]^M$	$x \in y$
$[\varphi \wedge \psi]^M$	$\varphi^M \wedge \varphi^M$
$[\neg\varphi]^M$	$\neg[\varphi^M]$
$[\exists_x \varphi]^M$	$\exists_x[x \in M \wedge \varphi^M]$

As noted earlier, every connective, as well as the universal quantifier, can be built up from $\wedge$, $\neg$, and $\exists$. Relativisations to M of formulas involving these connectives and quantifier can be constructed similarly.

For a sentence φ, if φ^M holds, we say φ is *true in* M. If S is a collection of sentences that are all true in M, we say M is a *model* for S. In particular, if the Zermelo-Fraenkel axioms are all true in M, then M is a model for Zermelo-Fraenkel set theory.

1.2 Axioms

In this section, we will recall the Zermelo-Fraenkel axioms for sets that will be used extensively throughout this thesis. Some of these axioms are redundant in the presence of others, but are used in different versions of Zermelo-Fraenkel set theory – for instance, the axioms of infinity and choice, which can be used to derive other axioms, are included in Zermelo's original 1908 introduction of his axioms [4], but excluded from his 1930 revision [18].

Axiom 1 (extensionality)**.** For all sets X and Y, if $x \in X \Leftrightarrow x \in Y$ for all x, then $X = Y$.

Axiom schema 2 (restricted comprehension)**.** If $\varphi(x, v_1, \ldots, v_n)$ is a formula with free variables $x, v_1, \ldots, v_n$, then for any set X and arbitrary sets $v_1, \ldots, v_n$, there exists a set

$$Y = \{x \in X \mid \varphi(x, v_1, \ldots, v_n)\}.$$

Axiom 3 (pairing)**.** If X and Y are sets, then there exists a set that has X and Y as elements.

For any sets X and Y, the axiom of pairing, together with the axiom schema of restricted comprehension, ensures that a set $\{X, Y\} = \{z \in Z \mid [z = X] \vee [z = Y]\}$ with exactly X and Y as elements exists, where Z is some set that has both X and Y as elements. Furthermore, by setting $Y = X$, we get that the singleton set $\{X\} = \{X, X\}$ exists for each X.

Since $\{X, Y\} = \{Y, X\}$, we need to add further structure if we want to capture the concept of 'order': for any sets X and Y, we define the *ordered pair* (X, Y) to be the set

$$(X, Y) = \{\{X\}, \{X, Y\}\}.$$

We can then define *ordered n-tuples* recursively, using (X, Y) as the base case:

$$(X_1, X_2, \ldots, X_{n+1}) = ((X_1, X_2, \ldots, X_n), X_{n+1}).$$

Ordered n-tuples are used to define the *cartesian product* of sets:

$$X_1 \times \ldots \times X_n = \{(x_1, \ldots, x_n) \mid \forall_{i \in \{1,\ldots,n\}}\, x_i \in X_i\}.$$

Axiom 4 (power set)**.** For each set X, there exists a set

$$\mathsf{P}X = \{Y \mid Y \subseteq X\},$$

called the *power set* of X.

Axiom 5 (union)**.** For every set X, there exists a set

$$\bigcup X = \{y \mid \exists_x [y \in x \wedge x \in X]\},$$

called the *union* of X.

One can also define the *intersection* of a set, which is a set by the axiom schema of restricted comprehension:

$$\bigcap X = \{y \in \bigcup X \mid \forall_{x \in X}\, y \in x\}.$$

This definition is noted in [19], although some authors prefer to define only a 'nonempty' intersection (to see why, notice that $\bigcap X$ becomes the class of all sets when the requirement $y \in \bigcup X$ is dropped). When taking the union and intersection of a two-element set $\{X, Y\}$, we may use the *binary union* notation $X \cup Y$ and the *binary intersection* notation $X \cap Y$. Notice that nothing prevents us from considering the union and intersection of a class of sets.

While the preceding axioms tell us how we can 'construct' sets using other sets, a set theory with no sets at all would be consistent with all of these axioms. Thus, if we want to ensure that there are any sets at all, we need to assert the existence of a set. The following two axioms do just that.

Axiom 6 (empty set)**.** There exists a set $\varnothing$ with no elements.

Axiom 7 (infinity)**.** There exists a set X such that

(a) $\varnothing \in X$;

(b) $x \cup \{x\} \in X$ for every $x \in X$.

An axiom of infinity is included in Zermelo's 1908 exposition of axiomatic set theory [4], although he requires that $\{x\} \in X$ (rather than $x \cup \{x\} \in X$) must hold for every

$x \in X$. He omits it in his 1930 revision of his earlier work [18], which allows one to consider a theory of finite sets.

It is possible to rephrase the axiom of infinity in such a way that it does not use the empty set in the definition of the infinite set X. Then, given the axiom schema of restricted comprehension, we can deduce the existence of the empty set:

$$\varnothing = \{x \in X \mid \neg[x = x]\}.$$

Axiom 8 (choice)**.** If X is a set of mutually disjoint sets, then there exists a subset of $\bigcup X$ that has exactly one element in common with each element of X.

While the axiom of choice was introduced in Zermelo's 1908 work [4], he leaves it out in his 1930 work [18], citing a difference in character to the other axioms. However, the axiom of choice is equivalent to the 'well-ordering principle', which asserts that each set is bijective to some ordinal number. We assume it in this thesis.

In 1922, Fraenkel found that Zermelo's axioms were insufficient to guarantee the existence of certain sets that were widely used, for instance any set that contains an infinite set and is closed under the power set operation (see e.g. [20]). He formulated the axiom schema of replacement for this purpose in [13].

Axiom schema 9 (replacement)**.** If $F = \{(x, y) \mid \varphi(x, y, v_1, \dots, v_n)\}$ is a class function with given parameters $v_1, \dots, v_n$ and X is a set, then FX is a set.

1.3 Ordered sets

In this section we recall some definitions concerning ordered sets, which we will use extensively in every part of the thesis. For instance, the logical connectives $\vee$ and $\wedge$ in Section 1.1 can be considered to define a lattice of (equivalence classes of) formulas (hence the overlapping notation), although we do not explore this idea. In Section 1.2, $(\mathrm{P}X, \subseteq)$ is a complete lattice for any set X, in which the join of $A \subseteq \mathrm{P}X$ is given by $\bigcup A$, and the meet of A is given by $\bigcap A$ when A is nonempty and X when A is empty. Building on Section 2.1, which examines properties of posets, ordinal systems are defined as special kinds of posets in Section 2.2, and are shown to be well-ordered in Theorem 50. For 'concrete ordinal systems' in particular, the order is given by the element relation $\in$. In Section 2.5, join semi-lattices are the structures used to define and characterise 'collecting systems'. In Chapter 3, the 'specialisation preorder' arising from a topology we define on a 'limit-successor system' (X, L, s) is a crucial ingredient in most of the proofs in that chapter, including the extensive proof of Theorem 78. This topology τ, like any 'Alexandrov' topology, is itself a complete lattice under the relation $\subseteq$, where the join of $A \subseteq \tau$ is given by $\bigcup A$, and the meet of A is given by $\bigcap A$ when A is nonempty and X when A is empty.

Definition 10 (preorder)**.** A *(non-strict) preorder* on a set X is a binary relation $\leqslant$ with the following properties.

- Reflexivity: $x \leqslant x$ for each $x \in X$.
- Transitivity: for any $x, y, z \in X$, if $x \leqslant y \leqslant z$, then $x \leqslant z$.

The ordered pair $(X, \leqslant)$ is called a *preordered set*.

We often do not distinguish between a preordered set $(X, \leqslant)$ and its underlying set X. We write $X = (X, \leqslant)$ to indicate this.

Definition 11 (equivalence relation, equivalence class and quotient set)**.** An *equivalence relation* on a set X (not usually thought of as an order relation) is a preorder $\sim$ with the following additional property.

- Symmetry: for any $x, y \in X$, if $x \sim y$, then $y \sim x$.

The *equivalence class* $[x]$ of an element $x \in X$ is then given by

$$[x] = \{y \mid y \sim x\},$$

and the *quotient set* $X/\!\sim$ of X by $\sim$ is given by

$$X/\!\sim\, = \{[x] \mid x \in X\}.$$

Definition 12 (partial order)**.** A *partial order* on a set X is a preorder $\leqslant$ with the following additional property.

- Antisymmetry: for any $x, y \in X$, if $x \leqslant y$ and $y \leqslant x$, then $x = y$.

The ordered pair $(X, \leqslant)$ is called a *partially ordered set* or a *poset*.

Notation 13 (strict order)**.** For any poset $(X, \leqslant)$, we write $<$ for the *strict* order relation corresponding to the *non-strict* order $\leqslant$:

$$[x < y] \quad \Leftrightarrow \quad [x \leqslant y] \wedge [x \neq y].$$

Definition 14 (minimal and maximal)**.** Let $(X, \leqslant)$ be a poset, and let $S \subseteq X$. An element $x \in S$ is said to be *minimal* if there is no $y \in S$ such that $y < x$, and *maximal* if there is no $y \in S$ such that $x < y$.

Definition 15 (minimum and maximum)**.** Let $(X, \leqslant)$ be a poset, and let $S \subseteq X$. An element $x \in S$ is called the *minimum* of S, denoted by $\min S$, if $x \leqslant y$ for all $y \in S$. If such an element exists, it is the unique minimal element of S. Likewise, $x \in S$ is called the *maximum* of S, denoted by $\max S$, if $y \leqslant x$ for all $y \in S$.

Note that a set S of elements of a poset can have at most one minimum element as well as at most one maximum element. Also, when a minimum/maximum element exists, it is necessarily the unique minimal/maximal element.

Definition 16 (total order)**.** A *total order* (or *linear order*) on a set X is a partial order with the following additional property.

- Totality: $x \leqslant y$ or $y \leqslant x$ for any $x, y \in X$.

The ordered pair $(X, \leqslant)$ is called a *totally ordered set*.

Definition 17 (well-order)**.** A *well-order* on a set X is a partial order $\leqslant$ with the following additional property.

- Well-foundedness: each nonempty subset of X has a minimum element.

The ordered pair $(X, \leqslant)$ is called a *well-ordered set*. Any well-ordered set is also totally ordered, since $\min\{x, y\}$ is defined for all $x, y \in X$.

Definition 18 (join and meet)**.** In a poset $(X, \subseteq)$, the *join* $\bigvee S$ and the *meet* $\bigwedge S$ of a subset S, if they are defined, are given by

$$\bigvee S = \min\{x \in X \mid \forall_{y \in S}[y \leqslant x]\},$$
$$\bigwedge S = \max\{x \in X \mid \forall_{y \in S}[x \leqslant y]\}.$$

For two elements $x, y \in X$, we also use the *binary join* and *binary meet* notation: respectively,

$$x \vee y = \bigvee\{x, y\} = \min\{z \in X \mid [x \leqslant z] \wedge [y \leqslant z]\},$$
$$x \wedge y = \bigwedge\{x, y\} = \max\{z \in X \mid [z \leqslant x] \wedge [z \leqslant y]\}.$$

Definition 19 (join and meet semilattices)**.** A *join semilattice* is a partially ordered set $(X, \leqslant)$ in which the binary join $x \vee y$ is defined for all $x, y \in X$. A *meet semilattice* is a partially ordered set $(X, \leqslant)$ in which the binary meet $x \wedge y$ is defined for all $x, y \in X$.

Definition 20 (lattice)**.** A *lattice* is a partially ordered set $(X, \leqslant)$ in which the binary join $x \vee y$ as well as the binary meet $x \wedge y$ are defined for all $x, y \in X$ (i.e. $(X, \leqslant)$ is a join semi-lattice as well as a meet semi-lattice). A *complete lattice* is a lattice in which the join $\bigvee S$ and the meet $\bigwedge S$ of every subset $S \subseteq X$ are defined.

Definition 21 (order isomorphism)**.** An *order isomorphism* between preordered sets $(X, \leqslant_X)$ and $(Y, \leqslant_Y)$ is a bijection $f\colon X \to Y$ such that for all $x, y \in X$,

$$x \leqslant_X y \quad \Leftrightarrow \quad f(x) \leqslant_Y f(y).$$

If an order isomorphism between X and Y exists, we call X and Y *order-isomorphic*.

It is worth noting that if $(X, \leqslant_X)$ and $(Y, \leqslant_Y)$ are posets, injectivity of f follows easily from antisymmetry of $\leqslant_X$.

1.4 Natural numbers

According to the approach of Richard Dedekind [2], a *natural number system* in set theory is defined as a triple $(\mathbb{N}, 0, s)$ – where $\mathbb{N}$ is a set, s is a function $\mathbb{N} \to \mathbb{N}$, and 0 is a distinguished element of $\mathbb{N}$ – that satisfies the following axioms.

(N1) 0 does not belong to the image of s.

(N2) s is injective.

(N3) $X = \mathbb{N}$ for any subset X of $\mathbb{N}$ that is closed under s and contains 0.

Axiom (N3) is known as the principle of *induction*, and a proof that uses it is called a *proof by induction*. Axioms (N1–3) are also known in the literature as Peano axioms [3]. Using these axioms, it is possible to establish the following 'universal property' of the natural number system. This property is also known as the principle of *recursive definition*, or simply *recursion*.

Theorem 22 (recursion)**.** *Let $(\mathbb{N}, 0, s)$ be a natural number system and let (X, x, t) be a triple such that X is a set, x is a distinguished element of X and t is a function $X \to X$. Then there exists a unique function $f\colon \mathbb{N} \to X$ that satisfies $f(0) = x$ and $f(s(n)) = t(f(n))$ for all $n \in \mathbb{N}$.*

The theorem above is originally due to Dedekind [2]. William Lawvere used the universal property identified in this theorem as the defining property for the natural number system in his 'elementary theory of the category of sets' (see [21]).

In John von Neumann's approach [14], natural numbers are defined as concrete sets. The version of the axiom of infinity given in Section 1.2 allows us to define the set ω of von Neumann natural numbers without further ado. Consider the formula

$$\varphi(X) = [\varnothing \in X] \wedge \forall_{x \in X}[x \cup \{x\} \in X].$$

By the axiom of infinity, there exists a set X that satisfies $\varphi(X)$. Then, by the axiom schema of restricted comprehension, we can define a set

$$\omega = \{x \in X \mid [x = \varnothing] \vee \forall_Y[\varphi(Y) \Rightarrow x \in Y]\},$$

i.e. the intersection of all sets Y satisfying $\varphi(Y)$. We call ω the set of *von Neumann natural numbers*. One can then prove that the triple $(\omega, \varnothing, t)$, where t is the function $\omega \to \omega$ that maps each $x \in \omega$ to $x \cup \{x\}$, is a natural number system.

The notion of a von Neumann natural number, as well as all the remaining notions recalled in this section, are standard in set theory (see e.g. [15, 22]).

If X and Y are sets and there exists a bijection $f\colon X \to Y$, we write $X \approx Y$ and say that X and Y are *bijective*. It is easy to show that $\approx$ is an equivalence relation, i.e. reflexive, transitive and symmetric (see Definition 10).

If a set X is bijective to a von Neumann natural number n, we say X is *finite*. If X is not finite, we say X is *infinite*. If X is bijective to ω, we say X is *countably infinite* (note that ω is infinite). If X is either finite or countably infinite we say X is *countable*. If X is not countable, we say X is *uncountable* or *uncountably infinite*. If we assume the axiom of choice, each set has a finite subset, and each infinite set has a countably infinite subset.

1.5 Universes

One way to incorporate discussions of classes as sets in axiomatic set theory is to consider a smaller 'universe of sets' within axiomatic set theory. Its elements will be the objects of study, but all classes of its elements will form sets rather than proper classes. Developing mathematics relative to a universe is typical in those subjects where sets of different sizes are needed; for instance, this is the approach followed in Saunders Mac Lane's exposition of category theory [5]. Universes will allow us to build a more general theory of ordinal numbers, in which the axioms of power set and infinity, for instance, need not hold. This section gives the definition and properties of the universes that we will use to develop our ordinal number systems.

Recall that a *transitive* set is a set X such that $x \subseteq X$ for all $x \in X$.

Definition 23 (universe)**.** A *universe* is a set $\mathfrak{U}$ satisfying the following axioms.

(U0) If $\mathfrak{U} \neq \varnothing$ then $\varnothing \in \mathfrak{U}$.

(U1) $\mathfrak{U}$ is a transitive set.

(U2) If $X, Y \in \mathfrak{U}$ then $\{X, Y\} \in \mathfrak{U}$.

(U3) $\bigcup fX \in \mathfrak{U}$ for any $X \in \mathfrak{U}$ and any function $f\colon X \to \mathfrak{U}$.

Note that this definition allows for the empty set $\varnothing$ to be a universe. Non-trivial examples can be found in standard set theory. In particular, for an 'infinite regular cardinal' κ, the set $H(\kappa)$ of sets that are 'hereditarily of cardinality less than κ', form a universe by Lemma 6.4 in Chapter IV of [12]. For instance, the set of 'hereditarily finite sets' is a universe – intuitively, hereditarily finite sets are finite sets which can be constructed from the empty set.

Remark 24 (Grothendieck universe)**.** A universe in the sense of Definition 23 is a universe in the sense of [6] (see also [23]) if it satisfies the following additional axiom:

(U4) If $X \in \mathfrak{U}$, then $\mathrm{P}X \in \mathfrak{U}$.

Remark 25. Note that (U4) makes (U0) and (U2) redundant. Since $\varnothing \in \mathrm{P}X$ for any set X, (U0) follows easily from (U1) and (U4). To see why (U2) follows, consider $X, Y \in \mathfrak{U}$. As we have just remarked, (U4) implies (U2), i.e. $\varnothing \in \mathfrak{U}$. Then by (U4), $\{\varnothing, \{\varnothing\}\} = \mathrm{PP}\varnothing \in \mathfrak{U}$. Furthermore, $\{X\} \in \mathrm{PP}X \subseteq \mathfrak{U}$ and $\{Y\} \in \mathrm{PP}Y \subseteq \mathfrak{U}$. Now define a function $f\colon \mathrm{PP}\varnothing \to \mathfrak{U}$ by

$$f(x) = \begin{cases} \{X\} & \text{if } x = \varnothing; \\ \{Y\} & \text{otherwise.} \end{cases}$$

Then $\{X, Y\} = \bigcup\{\{X\}, \{Y\}\} = \bigcup f\mathrm{PP}\varnothing \in \mathfrak{U}$ by (U3).

Definition 26 ($\mathfrak{U}$-small, $\mathfrak{U}$-moderate and $\mathfrak{U}$-large sets)**.** An element of a given universe is called a $\mathfrak{U}$-*small* set (analogous to a set). A subset of the universe that is not an element thereof is called a $\mathfrak{U}$-*moderate* set (analogous to a proper class). A set that is neither is called a $\mathfrak{U}$-*large* set (which has no analogue in classical Zermelo-Fraenkel set theory). Sets that are bijective to $\mathfrak{U}$-small and $\mathfrak{U}$-moderate are called *essentially* $\mathfrak{U}$-*small* and *essentially* $\mathfrak{U}$-*moderate*, respectively.

Theorem 27. *Any universe $\mathfrak{U}$ is closed under the following standard set-theoretic constructions.*

(a) *Singletons: if* $X \in \mathfrak{U}$*, then* $\{X\} \in \mathfrak{U}$*.*

(b) *Union: if* $X \in \mathfrak{U}$*, then* $\bigcup X \in \mathfrak{U}$*.*

(c) *Subsets: if* $X \subseteq Y \in \mathfrak{U}$*, then* $X \in \mathfrak{U}$*.*

(d) *Cartesian products (binary): if* $X, Y \in \mathfrak{U}$*, then* $X \times Y \in \mathfrak{U}$*.*

(e) *Disjoint union: if* $X \in \mathfrak{U}$*, then* $\sum X \in \mathfrak{U}$*, where*

$$\sum X = \bigcup \{x \times \{x\} \mid x \in X\}.$$

(f) *Quotient sets: if* $X \in \mathfrak{U}$ *and* $\sim$ *is an equivalence relation on* X*, then* $X/\!\sim\, \in \mathfrak{U}$*.*

(g) *Replacement: if* $X \in \mathfrak{U}$ *and* f *is a function* $X \to \mathfrak{U}$*, then* $fX \in \mathfrak{U}$*.*

Proof.

a. If $X \in \mathfrak{U}$, then $\{X\} = \{X, X\} \in \mathfrak{U}$ by (U2).

b. If $X \in \mathfrak{U}$, then $\bigcup X = \bigcup 1_X X \in \mathfrak{U}$ by (U3).

c. Let $X \subseteq Y$. If $Y \in \mathfrak{U}$ and $X = \varnothing$, then $X \in \mathfrak{U}$ by (U0). If $X \neq \varnothing$, then let $x \in X$. Consider the function $f\colon Y \to \mathfrak{U}$ defined by

$$f(y) = \begin{cases} \{y\} & \text{if } y \in X, \\ \{x\} & \text{if } y \notin X. \end{cases}$$

Then $X = \bigcup fY \in \mathfrak{U}$ by (U3).

d. Let $X, Y \in \mathfrak{U}$. Then

$$\{(x,y)\} = \{\{\{x,y\},\{x\}\}\} \in \mathfrak{U}$$

for each $x \in X$ and $y \in Y$ by (U2). For each $x \in X$, define a function $f_x\colon Y \to \mathfrak{U}$ by $f_x(y) = \{(x,y)\}$. Then $\{(x,y) \mid y \in Y\} = \bigcup f_x Y \in \mathfrak{U}$ for each $x \in X$ by (U3).

Now define a function $g\colon X \to \mathfrak{U}$ by $g(x) = \{\{(x,y) \mid y \in Y\}\}$. Then by (U3),

$$\begin{aligned} X \times Y &= \{(x,y) \mid x \in X \wedge y \in Y\} \\ &= \bigcup \{\{(x,y) \mid y \in Y\} \mid x \in X\} \\ &= \bigcup gX \in \mathfrak{U}. \end{aligned}$$

e. If $X \in \mathfrak{U}$, and $f\colon X \to \mathfrak{U}$ is defined by $f(x) = x \times \{x\}$, then by (U3),

$$\sum X = \bigcup\{x \times \{x\} \mid x \in X\} = \bigcup fX \in \mathfrak{U}.$$

f. Let $X \in \mathfrak{U}$, let $\sim$ be an equivalence relation on X (see Definition 11), and let q be the function $q\colon X \to \mathfrak{U}$ defined by $q(x) = \{[x]\}$. Then $X/\!\sim\, = \bigcup qX \in \mathfrak{U}$ by (U3).

g. Let $X \in \mathfrak{U}$ and let f be a function $X \to \mathfrak{U}$. By (a), we can define a function $g\colon \mathfrak{U} \to \mathfrak{U}$ such that $g(y) = \{y\}$ for each $y \in \mathfrak{U}$. Then $g \circ f$ is a function $X \to \mathfrak{U}$ such that for all $x \in X$,

$$(g \circ f)(x) = \{f(x)\}.$$

The image of X under $g \circ f$ is then

$$(g \circ f)X = \{\{f(x)\} \mid x \in X\}.$$

Then by (U3),

$$\begin{aligned} fX &= \{f(x) \mid x \in X\} \\ &= \bigcup\{\{f(x)\} \mid x \in X\} \\ &= \bigcup(g \circ f)X \in \mathfrak{U}. \end{aligned}$$

■

Remark 28. Theorem 27 is significant because it demonstrates the true strength of a universe. Since the elements of a universe are all sets, $\mathfrak{U}$ 'inherits' the axioms of extensionality and choice from the Zermelo-Fraenkel set theory in which $\mathfrak{U}$ lies – that is, relativisations (see Section 1.1) of those axioms to $\mathfrak{U}$ hold. By (U2), a relativisation of the axiom of pairing to $\mathfrak{U}$ holds, and by Theorem 27, relativisations to $\mathfrak{U}$ of the axiom of union and the axiom schemas of restricted comprehension and replacement also hold. This means that $\mathfrak{U}$ is a model for these axioms. If $\mathfrak{U}$ additionally satisfies (U4), then a relativisation of the axiom of power set to $\mathfrak{U}$ also holds, which makes $\mathfrak{U}$ a model of Zermelo-Fraenkel set theory (without the axiom of infinity).

Lemma 29. *If a universe $\mathfrak{U}$ is nonempty, then $\omega \subseteq \mathfrak{U}$. If $\mathfrak{U}$ has at least one infinite element, then $\omega \in \mathfrak{U}$.*

Proof. Recall that ω is a natural number system where the successor of an element $n \in \omega$ is given by $n \cup \{n\}$. To see that $\omega \subseteq \mathfrak{U}$, notice that $0 = \varnothing \in \omega$ by (U0), and that $n + 1 = n \cup \{n\} \in \mathfrak{U}$ whenever $n \in \mathfrak{U}$, for all $n \in \omega$ (by (U2) and Theorem 27 (a–b)). We can conclude by induction that $\omega \subseteq \mathfrak{U}$.

Now suppose $\mathfrak{U}$ has at least one infinite element X. Then X has a countably infinite subset $Y \in \mathfrak{U}$ by Theorem 27 (c) and the axiom of choice. Then a bijection $f\colon Y \to \omega$ exists, and since $\omega \subseteq \mathfrak{U}$, this gives us $fY = \omega \in \mathfrak{U}$ by Theorem 27 (g). ∎

Definition 30 (restricted power set)**.** The *restricted power set* of a set X relative to a universe $\mathfrak{U}$ is the set

$$\mathsf{P}_{\mathfrak{U}}X = \{A \subseteq X \mid \exists_{B\in\mathfrak{U}}[A \approx B]\},$$

i.e. the set of essentially $\mathfrak{U}$-small subsets of X.

Lemma 31. *When $\mathfrak{U}$ is nonempty, for any set X and its finite subset $Y \subseteq X$, we have $Y \in \mathsf{P}_{\mathfrak{U}}X$.*

Proof. Since $\omega \subseteq \mathfrak{U}$ by Lemma 29, this follows from the fact that a finite set is defined as a set bijective to a von Neumann natural number. ∎

Lemma 32. *If $A \subseteq B$ and $B \in \mathsf{P}_{\mathfrak{U}}X$, then $A \in \mathsf{P}_{\mathfrak{U}}X$.*

Proof. If $A \subseteq B \approx C \in \mathfrak{U}$, then A is bijective to a subset of C. ∎

Lemma 33. *If $C \in \mathsf{P}_{\mathfrak{U}}\mathsf{P}_{\mathfrak{U}}X$ then $\bigcup C \in \mathsf{P}_{\mathfrak{U}}X$. In particular, this implies that if $I \in \mathsf{P}_{\mathfrak{U}}X$, then for any function $f\colon I \to \mathsf{P}_{\mathfrak{U}}X$, we have $\bigcup fI \in \mathsf{P}_{\mathfrak{U}}X$.*

Proof. Suppose $C \in \mathsf{P}_{\mathfrak{U}}\mathsf{P}_{\mathfrak{U}}X$. Then there is a bijection $h\colon C' \to C$, where $C' \in \mathfrak{U}$. Since for each element $c \in C$ we have $c \in \mathsf{P}_{\mathfrak{U}}X$, by axiom of choice we have a function $g\colon C \to \mathfrak{U}$ such that $c \approx g(c)$ for each $c \in C$. Let g_c denote a bijection $g_c\colon c \to g(c)$ (we again use the axiom of choice to select such a bijection for each $c \in C$). Now, define a function $k\colon C' \to \mathfrak{U}$ as follows:

$$k(c') = \{(x, c') \mid x \in g(h(c'))\}.$$

Then $\bigcup kC' \in \mathfrak{U}$ by (U3). Consider the function $f\colon \bigcup kC' \to \bigcup C$ defined by

$$f(x, c') = g^{-1}_{h(c')}(x).$$

This function is a surjection. Indeed, for each $y \in \bigcup C$ there is a $c \in C$ such that $y \in c$ and so,

$$\begin{aligned} f(g_c(y), h^{-1}(c)) &= g^{-1}_{h(h^{-1}(c))}(g_c(y)) \\ &= g_c^{-1}(g_c(y)) \\ &= y. \end{aligned}$$

So $\bigcup C$ is bijective to a quotient set of $\bigcup kC'$, and thus $\bigcup C \in \mathsf{P}_{\mathfrak{U}}X$. This proves the first part of the lemma.

Now suppose $I \in \mathsf{P}_{\mathfrak{U}}X$ and let $f\colon I \to \mathsf{P}_{\mathfrak{U}}X$ be a function. Then $fI \in \mathsf{P}_{\mathfrak{U}}\mathsf{P}_{\mathfrak{U}}X$. By what we have just proved, $\bigcup fI \in \mathsf{P}_{\mathfrak{U}}X$. ∎

Lemma 34. *Given a function $f\colon X \to Y$,*

$$A \in \mathsf{P}_{\mathfrak{U}}X \quad \Rightarrow \quad fA \in \mathsf{P}_{\mathfrak{U}}Y.$$

Proof. If $A \approx A' \in \mathfrak{U}$, then fA is bijective to a suitable quotient of A'. ∎

Lemma 35. *If $A \in \mathsf{P}_{\mathfrak{U}}X$ and $B \in \mathsf{P}_{\mathfrak{U}}Y$, then $A \times B \in \mathsf{P}_{\mathfrak{U}}(X \times Y)$.*

Proof. Let $A \approx A' \in \mathfrak{U}$ and $B \approx B' \in \mathfrak{U}$. Since $\mathfrak{U}$ is closed under cartesian products (Theorem 27 (d)), we have $A \times B \approx A' \times B' \in \mathfrak{U}$, giving us $A \times B \in \mathsf{P}_{\mathfrak{U}}(X \times Y)$. ∎

Lemma 36. $\mathsf{P}_{\mathfrak{U}}\mathfrak{U} = \mathfrak{U}$.

Proof. Let $A \in \mathsf{P}_{\mathfrak{U}}\mathfrak{U}$, i.e. $A \subseteq \mathfrak{U}$ such that $A \approx B \in \mathfrak{U}$. Then a bijection $f\colon B \to A$ exists, and since $A \subseteq \mathfrak{U}$, this gives us $A = fB \in \mathfrak{U}$ by Theorem 27 (g).

Now consider $A \in \mathfrak{U}$. Then $A \subseteq \mathfrak{U}$ by (U1), and since $A \approx A \in \mathfrak{U}$, this gives us $A \in \mathsf{P}_{\mathfrak{U}}\mathfrak{U}$. We can conclude that $\mathsf{P}_{\mathfrak{U}}\mathfrak{U} = \mathfrak{U}$. ∎

Chapter 2

Ordinal systems

In this chapter we enter the central theme of this thesis: an axiomatic formulation of ordinal systems as abstract objects, and an exploration of their properties.

In Section 2.1 we define some functions that we will need for defining our ordinal systems. Crucially, this includes the 'incremented join' function $\overset{+}{\bigvee}$, which assigns to a subset of a poset the smallest element of the poset that is an upper bound for the subset, but not an element of it. The incremented join captures both the successor and the limit properties of ordinal numbers – the successor of an ordinal is computed as $x^+ = \overset{+}{\bigvee}\{x\}$, while limit ordinals are incremented joins of sets of ordinals having no maximum element.

In Section 2.2, we use the incremented join function to define and characterise ordinal systems relative to a universe $\mathfrak{U}$. Section 2.3 formulates and proves transfinite induction and recursion for these ordinal systems.

In Section 2.4, we prove that the set of concrete (von Neumann) ordinals that are elements of $\mathfrak{U}$ is a special case of an ordinal system relative to $\mathfrak{U}$.

In Section 2.5, we formulate and prove the first of the two universal properties of ordinal systems obtained in this thesis, which presents them as initial objects in certain categories of ordered sets.

2.1 Foundation

The definitions and lemmas that follow will provide the foundation necessary for the development of our ordinal systems. Since these definitions and lemmas all apply to partially ordered sets (see Definition 12), they are stated for an arbitrarily fixed poset $(X, \leqslant)$ throughout this section.

Definition 37 (upper and lower complements)**.** For a subset S of X, we define the *upper complement* $S^<$ of S as the set

$$S^< = \{x \in X \mid \forall_{y \in S}[y < x]\}.$$

Likewise, we define the *lower complement* $S^>$ of S as

$$S^> = \{x \in X \mid \forall_{y \in S}[x < y]\}.$$

The following lemma follows from a general well-known fact that any relation between sets induces a Galois connection between the power sets of those sets (see e.g. [24]). We present the proof of this particular case for completeness.

Lemma 38. *For subsets S of X, the mappings $S \mapsto S^<$ and $S \mapsto S^>$ form a Galois connection from $\mathsf{P}X$ to itself induced by the relation $<$, i.e.*

$$S \subseteq T^> \;\Leftrightarrow\; T \subseteq S^<.$$

Proof. For the forwards implication, suppose $S \subseteq T^>$. Consider any $x \in T$. Then $y < x$ for all $y \in S$, which implies $x \in S^<$. This proves $T \subseteq S^<$. Applying this result to the dual poset $(X, \geqslant)$, we get the backwards implication (the upper/lower complement in the dual poset is the lower/upper complement in the original poset). ■

Lemma 39. *The following hold for all subsets S of X:*

(a) $S \cap S^> = S \cap S^< = \varnothing$;

(b) $S \subseteq S^{><}$ *and* $S \subseteq S^{<>}$.

Proof. We prove this as follows.

a. Consider any $x \in S$. Since $x \not< x$, neither $\forall_{y \in S}[x < y]$ nor $\forall_{y \in S}[y < x]$ holds, and thus $x \notin S^>$ and $x \notin S^<$. Since this is true for all $x \in S$, we have $S \cap S^> = S \cap S^< = \varnothing$.

b. Of course $S^> \subseteq S^>$. Then by Lemma 38, $S \subseteq S^{><}$. Likewise, since $S^< \subseteq S^<$, Lemma 38 gives us $S \subseteq S^{<>}$. ■

Definition 40 (incremented join)**.** We define the *incremented join* $\overset{+}{\bigvee} S$ of a subset S of X (when it exists) as follows:

$$\overset{+}{\bigvee} S = \min S^<.$$

Remark 41. Note that since $S \cap S^< = \varnothing$ (Lemma 39), the incremented join of S is never an element of S.

Definition 42 (successor)**.** We define the *successor* x^+ (if it exists) of an element x of X as follows:

$$x^+ = \curlyvee\{x\}.$$

If all elements of a subset S of X have successors, we write

$$S^+ = \{x^+ \mid x \in S\}$$

for the image of S under the successor relation.

Remark 43. The successor x^+ of x is the smallest element larger than x, which aligns with the usual idea of a successor.

Lemma 44. *For any $x, y \in X$ such that x^+ is defined, we have:*

(L1) $x < x^+$;

(L2) *there is no z such that $x < z < x^+$;*

(L3) $x < y \;\Leftrightarrow\; x^+ \leqslant y$.

When, furthermore, $(X, \leqslant)$ is a totally ordered set, and y^+ is defined in (L5–7), we also have the following:

(L4) $x = \curlyvee\{x\}^>$;

(L5) $x < y^+ \;\Leftrightarrow\; x \leqslant y$;

(L6) $x^+ = y^+ \;\Leftrightarrow\; x = y$;

(L7) $x < y \;\Leftrightarrow\; x^+ < y^+$.

Proof. Consider $x, y \in X$ such that x^+ is defined. (L1–2) follow trivially from the definition of x^+ as the smallest element of X that is greater than x.

L3. If $x^+ \leqslant y$, then $x < y$ follows from (L1). Conversely, if $x < y$ then $y \in \{x\}^<$, and so $x^+ \leqslant y$ (recall that $x^+ = \min\{x\}^<$).

Now suppose $(X, \leqslant)$ is totally ordered (see Definition 16), and, for (L5–7), that y^+ is defined.

L4. Since $(X, \leqslant)$ is totally ordered and $\{x\}^> \cap \{x\}^{><} = \varnothing$ (Lemma 39), it holds for all $y \in \{x\}^{><}$ that $x \leqslant y$. By Lemma 39, $x \in \{x\} \subseteq \{x\}^{><}$, and it follows that $x = \min\{x\}^{><} = \curlyvee\{x\}^>$.

L5. Suppose $x \leqslant y$. Since $y < y^+$ by (L1), this gives us $x < y^+$. Now suppose $x < y^+$. Since we know by (L2) that it cannot hold that $y < x < y^+$, we can conclude that $x \leqslant y$.

L6. Suppose $x^+ = y^+$. Since $x < x^+$ and $y < y^+$, and since, by (L2), it cannot hold that $x < y < x^+$ or $y < x < y^+$, we have $x \leqslant y$ and $y \leqslant x$, and thus $x = y$. The other direction is trivial.

L7. By (L3,5), $x < y \;\Leftrightarrow\; x^+ \leqslant y \;\Leftrightarrow\; x^+ < y^+$. ■

Lemma 45. *For any $S \subseteq X$, if x^+ exists for all $x \in S$, then $\bigvee S^+$ exists if and only if $\bar{\bigvee} S$ exists, and when they exist,*

$$\bigvee S^+ = \bar{\bigvee} S.$$

Proof. Consider a subset S of X such that x^+ exists for all $x \in S$. By (L3),

$$\{x \in X \mid \forall_{y \in S}[y^+ \leqslant x]\} = \{x \in X \mid \forall_{y \in S}[y < x]\},$$

and thus

$$\begin{aligned} \bigvee S^+ &= \min\{x \in X \mid \forall_{y \in S^+}[y \leqslant x]\} \\ &= \min\{x \in X \mid \forall_{y \in S}[y^+ \leqslant x]\} \end{aligned}$$

exists if and only if

$$\bar{\bigvee} S = \min\{x \in X \mid \forall_{y \in S}[y < x]\}$$

exists, and when they exist, they are equal. ■

Lemma 46. *Let $S \subseteq X$.*

(a) *If S does not have a largest element, then $\bigvee S$ exists if and only if $\bar{\bigvee} S$ exists, and when they exist, $\bigvee S = \bar{\bigvee} S$. Conversely, if $\bigvee S = \bar{\bigvee} S$, then S does not have a largest element.*

(b) *If S has a largest element $x = \max S$, then $\bar{\bigvee} S$ exists if and only if x^+ exists, and when they exist, $\bar{\bigvee} S = x^+$. Conversely, when $\leqslant$ is a total order, if $\bar{\bigvee} S = x^+$, then $x = \max S$.*

(c) *If $\bar{\bigvee} S^>$ exists, then $\bar{\bigvee} S^> = \min S$. Conversely, when $\leqslant$ is a total order, if $\min S$ exists, then $\min S = \bar{\bigvee} S^>$.*

Proof.

a. If S has no largest element, then

$$\{x \in X \mid \forall_{y \in S}[y \leqslant x]\} = \{x \in X \mid \forall_{y \in S}[y < x]\}$$

and so

$$\bigvee S = \min\{x \in X \mid \forall_{y \in S}[y \leqslant x]\}$$

exists if and only if

$$\dot{\bigvee} S = \min\{x \in X \mid \forall_{y \in S}[y < x]\}$$

exists, and when they exist, they are equal.

Now suppose instead that $\bigvee S = \dot{\bigvee} S$. Since $\dot{\bigvee} S$ is never an element of S (Remark 41), we conclude that S does not have a largest element.

b. Let $x = \max S$. Then $\{x\}^{<} = S^{<}$, and so $x^{+} = \min\{x\}^{<}$ exists if and only if $\dot{\bigvee} S = \min S^{<}$ exists, and when they exist, they are equal.

Now suppose, in the case of total order, that $\dot{\bigvee} S = x^{+}$ for some $x \in X$. Then S can only have elements that are strictly smaller than x^{+}, and thus each element of S is less than or equal to x by (L5). Also, since $x^{+} = \dot{\bigvee} S = \min S^{<}$, it does not hold that $x \in S^{<}$, and thus x is not strictly larger than every element of S. By total order, this implies that x is less or equal to one of the elements of S. We can then conclude that $x = \max S$.

c. Suppose $\dot{\bigvee} S^{>}$ exists. Since $S \subseteq S^{><}$ (Lemma 39), we must have $\dot{\bigvee} S^{>} \leqslant x$ for each $x \in S$. This, together with the fact that $\dot{\bigvee} S^{>}$ cannot be an element of $S^{>}$ (Remark 41), forces $\dot{\bigvee} S^{>}$ to be an element of S. Hence $\dot{\bigvee} S^{>} = \min S$. Suppose instead that $\leqslant$ is a total order and $\min S$ exists. Since again $S \subseteq S^{><}$, we get $\min S \in S^{><}$. Consider an element $x \in S^{><}$. Then x is not an element of $S^{>}$ and so we cannot have $x < \min S$. Therefore, $\min S \leqslant x$. This proves $\min S = \dot{\bigvee} S^{>}$. ■

2.2 Abstract ordinal systems

Definition 47 (ordinal system). An *ordinal system* relative to a universe $\mathfrak{U}$ is a partially ordered set $\mathcal{O} = (\mathcal{O}, \leqslant)$ satisfying the following axioms.

(O1) For all $X \subseteq \mathcal{O}$, if $X \neq \varnothing$, then $X^{>} \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$.

(O2) $\dot{\bigvee} X$ exists for each $X \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$.

We refer to elements of $\mathcal{O}$ as *ordinals*.

Remark 48. Note that if $\mathcal{O}$ is nonempty, then Axiom (O1) forces the universe $\mathfrak{U}$ to be nonempty as well. So for any ordinal $x \in \mathcal{O}$, we have $\{x\} \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$ (Lemma 31). Also note that Axiom (O1) is equivalent to its weaker form:

(O1′) $\{x\}^{>} \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$ for all $x \in \mathcal{O}$.

The equivalence does not require (O2). It is easy to see that (O1) $\Rightarrow$ (O1′). For the other direction, consider a nonempty $X \subseteq \mathcal{O}$. Then, picking any $x \in X$, we have $X^{>} \subseteq \{x\}^{>} \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$, and by Lemma 32, this means $X^{>} \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$.

Definition 49 (successor function, successor and limit ordinals)**.** Axiom (O2) implies that the mapping $x \mapsto x^{+}$ is a function $\mathcal{O} \to \mathcal{O}$. We call this function the *successor function* of the ordinal system $\mathcal{O}$. For each $x \in \mathcal{O}$, an element of $\mathcal{O}$ that has the form x^{+} is called a *successor ordinal* and the *successor* of x. An ordinal that is not a successor ordinal is called a *limit ordinal*.

Theorem 50. *Given a poset $(\mathcal{O}, \leqslant)$, the following claims are equivalent:*

(a) *$\mathcal{O}$ is an ordinal system relative to a universe $\mathfrak{U}$;*

(b) *$\mathcal{O}$ is a well-ordered set (see Definition 17) satisfying (O1′), and $X^{<} \neq \varnothing$ for all $X \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$.*

Proof. To see that (a) $\Rightarrow$ (b), consider an ordinal system $\mathcal{O}$ relative to a universe $\mathfrak{U}$. Let X be a nonempty subset of $\mathcal{O}$. Then $X^{>} \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$ by (O1), and thus $\bigvee\!\!\!\!/\, X^{>}$ exists by (O2). By Lemma 46 (c), $\bigvee\!\!\!\!/\, X^{>} = \min X$.

(b) $\Rightarrow$ (a) follows trivially from the fact that the nonempty set $X^{<}$ has a minimum element $\bigvee\!\!\!\!/\, X$ in the well-ordered set $\mathcal{O}$. ■

Remark 51. One of the consequences of this theorem is that each nonempty ordinal system has a smallest element. We denote this element by 0. Note that 0 is a limit ordinal. Since, by the same theorem, an ordinal system is a total order, the properties (L4–7) apply to an ordinal system. In particular, (L6) guarantees that the successor function is injective.

Lemma 52. *In an ordinal system, for an ordinal x the following conditions are equivalent:*

(a) *x is a limit ordinal;*

(b) *$\{x\}^{>}$ is closed under successors;*

(c) $x = \bigvee\{x\}^{>}$.

Proof. We have $x = \bigvee\!\!\!\!\forall\{x\}^{>}$ for any ordinal x by (L4). If x is a limit ordinal, then for each $y < x$ we have $y^{+} < x$ by (L3). So (a) $\Rightarrow$ (b).

If (b) holds, by (L1), we get that $\{x\}^{>}$ does not have a largest element. So $\bigvee\!\!\!\!\forall\{x\}^{>} = \bigvee\{x\}^{>}$ (Lemma 46). This gives (b) $\Rightarrow$ (c).

Suppose now that $x = \bigvee\{x\}^{>}$. If x were a successor ordinal $x = y^{+}$, then by (L5), y would be the join of $\{x\}^{>}$. However, $x \neq y$ by (L1), and therefore x must be a limit ordinal. Thus, (c) $\Rightarrow$ (a). ∎

The following theorem gives yet another way of thinking about an ordinal system.

Theorem 53. *A poset $(\mathcal{O}, \leqslant)$ is an ordinal system relative to a universe $\mathfrak{U}$ if and only if (O1) holds along with the following axioms.*

(O2a) *For all $X \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$, the join $\bigvee X$ exists.*

(O2b) *x^{+} exists for each $x \in \mathcal{O}$.*

Proof. This follows easily from (a–b) in Lemma 46. ∎

2.3 Transfinite induction and recursion

Transfinite induction and recursion are well known for well-ordered sets (see e.g. [19]), and ordinal systems are special kinds of well-ordered sets. Nevertheless, we formulate and prove these results in our context for the sake of completeness.

Theorem 54 (transfinite induction)**.** *Let $\mathcal{O}$ be an ordinal system and let $X \subseteq \mathcal{O}$ satisfy the following conditions.*

(I1) $X^{+} \subseteq X$.

(I2) *For every limit ordinal x, if $\{x\}^{>} \subseteq X$ then $x \in X$.*

Then $X = \mathcal{O}$.

Proof. Since $\mathcal{O}$ is well-ordered, if $\mathcal{O} \setminus X$ is nonempty, then it has a smallest element y. By (I1), y cannot be a successor of any $z < y$ in X. By (I2), it also cannot be a limit ordinal. This is a contradiction, since a limit ordinal is defined as one that is not a successor ordinal. ∎

We will also use the following alternative form of transfinite induction, which is obtained by 'strengthening' (I1) to match (I2).

Theorem 55 (strong transfinite induction). *Let $\mathcal{O}$ be an ordinal system and let $X \subseteq \mathcal{O}$ satisfy the following condition.*

(I′) *For every ordinal x, if $\{x\}^{>} \subseteq X$, then $x \in X$.*

Then $X = \mathcal{O}$.

Proof. If $\mathcal{O}$ is well-ordered, if $\mathcal{O} \setminus X$ is nonempty, then it has a smallest element y. But then $\{y\}^{>} \subseteq X$. This contradicts (I′). ∎

For natural numbers, the induction axiom (N3) states that any 'inductive' subset of $\mathbb{N}$ must equal $\mathbb{N}$. The following corollary to Theorem 54 captures a similar property of ordinal systems.

Corollary 56. *Let $\mathcal{O}$ be an ordinal system relative to a universe $\mathfrak{U}$, and let $X \subseteq \mathcal{O}$ satisfy the following conditions.*

(S1) *X is down-closed in $\mathcal{O}$, i.e. if $x < y \in X$ then $x \in X$, for all $x, y \in \mathcal{O}$.*

(S2) *X is an ordinal system relative to $\mathfrak{U}$ under the restriction of the order of $\mathcal{O}$.*

Then $X = \mathcal{O}$.

Proof. Consider any $x \in X$. Since x must have a successor in X by (S2), there exists $y \in X$ such that $x < y$. By (L3), it must hold that $x^+ \leqslant y$, where x^+ is the successor of x in $\mathcal{O}$. Since x is down-closed in $\mathcal{O}$, it also holds that $x^+ \in X$, and thus x^+ is thus also the successor of x in X, which means that (I1) holds.

Now consider any limit ordinal $x \in \mathcal{O}$. Suppose $\{x\}^{>} \subseteq X$. Then, since X is an ordinal system, $\{x\}^{>}$ must have an incremented join y in X. Since y must be bigger than each element of $\{x\}^{>}$, we see that y is in the set $\{x\}^{><}$, which has $x = \forall\{x\}^{>}$ as its minimum in $\mathcal{O}$. Since $x \leqslant y$, and X is down-closed in $\mathcal{O}$, it also holds that $x \in X$, which means that (I2) holds. Thus, by Theorem 54, $X = \mathcal{O}$. ∎

Theorem 57 (transfinite recursion). *Let $\mathcal{O}$ be an ordinal system, and let $X = (X, L, s)$, where X is a set, s is a function $X \to X$, and L is a function $\mathsf{P}_{\mathfrak{U}}X \to X$. Then there exists a unique function $f\colon \mathcal{O} \to X$ that satisfies the following conditions.*

(R1) *$f(x^+) = s(f(x))$ for any ordinal x.*

(R2) *$f(x) = L(\{f(y) \mid y < x\})$ for any limit ordinal x.*

Proof. For each ordinal $x \in \mathcal{O}$, let F_x be the set consisting of all functions

$$f_x \colon \{y \in \mathcal{O} \mid y \leqslant x\} \to X$$

that satisfy the following conditions:

(a) $f_x(y^+) = s(f(y))$ for any successor ordinal $y^+ \leqslant x$;

(b) $f_x(y) = L(\{f(z) \mid z < y\})$ for any limit ordinal $y \leqslant x$.

Note that any function f satisfying (R1–2) must have a subfunction in each set F_x. Also, for any $x \in \mathcal{O}$, a function $f_x \in F_x$ must have a subfunction in each set F_y where $y < x$.

We prove by transfinite induction that for each $x \in \mathcal{O}$, the set F_x contains exactly one function f_x.

Successor case: Suppose that for some ordinal x and each $y \leqslant x$ there exists a unique function $f_y \in F_y$. Then $g = f_x \cup \{(x^+, s(f_x(x)))\}$ satisfies (a–b), and thus $g \in F_{x^+}$.

Now consider any function $g' \in F_{x^+}$. Since g and g' must each have the unique function $f_x \in F_x$ as a subfunction, g and g' are identical on the domain $\{y \mid y \leqslant x\}$. However, since g and g' both satisfy condition (a), we get that $g'(x^+) = s(f_x(x)) = g(x^+)$ and thus $g' = g$ is the unique function in F_{x^+}.

Limit case: Suppose that for some limit ordinal x and each $y < x$ there exists a unique function $f_y \in F_y$. Then, for any two ordinals $y < y' < x$, the function f_y is a subfunction of $f_{y'}$, which is in turn a subfunction of every function in F_x. The relation

$$g = \bigcup\{f_y \mid y < x\} \cup \{(x, L(\{f_y(y) \mid y < x\}))\}$$

is then a function over the domain $\{y \mid y \leqslant x\}$ and, moreover, it is easy to see that $g \in F_x$.

Now consider any function $g' \in F_x$. Since for each ordinal $y < x$ the unique function $f_y \in F_y$ is a subfunction of both g and g', we see that they are identical on the domain $\{y \mid y < x\}$, and since they both satisfy condition (b), the following holds:

$$g(x) = L(\{g(y) \mid y < x\}) = L(\{g'(y) \mid y < x\}) = g'(x).$$

We can conclude that $g' = g$ is the unique function in F_x.

We showed that for each ordinal x, exactly one function $f_x \in F_x$ exists. Construct a function $f \colon \mathcal{O} \to X$ as follows: for each ordinal x,

$$f(x) = f_x(x).$$

It satisfies (R1–2) because each f_x satisfies (a–b). Since any function satisfying (R1–2) must have a subfunction in each set F_x, we can conclude that f is the unique function satisfying (R1–2). ■

2.4 Concrete ordinal systems

This section investigates the relationship between our abstract ordinal systems and the concrete system of von Neumann ordinals that is standard in the literature.

Definition 58 ($\mathfrak{U}$-ordinal)**.** For a universe $\mathfrak{U}$, define a $\mathfrak{U}$-*ordinal* to be a transitive set x that belongs to the universe $\mathfrak{U}$, is a well-ordered set under the relation

$$y \mathrel{\underline{\in}} z \quad \Leftrightarrow \quad [y \in z] \vee [y = z],$$

and has no element z that has the property $z \in z$ (in other words, $\in$ is the strict ordering for the partial order $\underline{\in}$).

Thus, a $\mathfrak{U}$-ordinal is nothing but a usual von Neumann ordinal number (see [25]) that happens to be an element of $\mathfrak{U}$. In other words, it is a von Neumann ordinal number 'internal' to the universe $\mathfrak{U}$. Thus, at least one direction in the following theorem is well known. We include a full proof for completeness.

Theorem 59. *A set O is the set of all $\mathfrak{U}$-ordinals if and only if the following conditions hold:*

(a) *$O \subseteq \mathfrak{U}$, and if $\varnothing \in \mathfrak{U}$, then $\varnothing \in O$;*

(b) *O is a transitive set;*

(c) *O is an ordinal system relative to $\mathfrak{U}$ under the relation $\underline{\in}$, with the corresponding strict ordering given by $\in$.*

When these conditions hold, the incremented join of $X \in \mathsf{P}_{\mathfrak{U}}O$ is given by $\curlyvee X = X \cup \bigcup X$.

Proof. We prove both directions. The proof of the last statement in the theorem is included in Step 1 (c).

Step 1: *We prove that if O is the set of all $\mathfrak{U}$-ordinals, then* (a–c) *hold. While proving* (c)*, we show that the incremented join of a set of $\mathfrak{U}$-ordinals $X \in \mathsf{P}_{\mathfrak{U}}O$ is given by $\curlyvee X = X \cup \bigcup X$.*

Let O be the set of all $\mathfrak{U}$-ordinals.

a. This follows easily from the definition of a $\mathfrak{U}$-ordinal as a transitive element of $\mathfrak{U}$.

b. To prove transitivity, let $y \in x \in O$. We want to show that y is a $\mathfrak{U}$-ordinal (i.e. an element of $\mathfrak{U}$ that is transitive, well-ordered by $\underline{\in}$, and has no element z such that $z \in z$). Since $x \in \mathfrak{U}$, we have $y \in \mathfrak{U}$ by (U1). Since y is a subset of the well-ordered set x, y must also be well-ordered, and since $z \in z$ holds for no element $z \in x$, it also holds for no element $z \in y$.

Next, we must prove that y is transitive. Let $z \in y$. Then $z \in x$, since x is transitive. We want to show that $z \subseteq y$, so suppose $t \in z$. Then $t \in x$, by transitivity of x. By the fact that $\underline{\in}$ is a total order on x, we must have $t \,\underline{\in}\, y$. We cannot have $t = y$, since $t \in z \in y$, and so $t \in y$.

c. We use the characterisation of an ordinal system given in Theorem 50 (b) as a well-ordered set satisfying (O1′), in which $X^< \neq \varnothing$ holds for all $X \in \mathsf{P}_{\mathfrak{U}}O$.

Notice that the relation $\underline{\in}$ is a partial order on O: reflexivity is obvious, while transitivity follows from each element of O being a transitive set. Furthermore, since $z \in z$ holds for no $\mathfrak{U}$-ordinal z, we have that $\in$ is the strict ordering on O corresponding to $\underline{\in}$.

We can easily prove, as follows, that O satisfies (O1′).

O1′. Consider $x \in O$. Then we have:

$$\{x\}^> = \{y \in O \mid y \in x\} \subseteq x.$$

Since x is a $\mathfrak{U}$-ordinal, $x \in \mathsf{P}_{\mathfrak{U}}O$ and so $\{x\}^> \in \mathsf{P}_{\mathfrak{U}}O$ (Lemma 32). Note that by (b) we actually have $x = \{x\}^>$, although we did not need this to establish (O1′).

Before we prove that O is well-ordered, we will first establish the following two facts.

Fact 1: *$x \cap y = \min(y \setminus x)$ for any two $\mathfrak{U}$-ordinals x and y such that $y \setminus x \neq \varnothing$.*

Let $a \in x \cap y$. Then $a \notin y \setminus x$, and hence $a \neq \min(y \setminus x)$. By the well-ordering of y, we then get that either $a \in \min(y \setminus x)$ or $\min(y \setminus x) \in a$. By transitivity of x and the fact that $\min(y \setminus x) \notin x$, the second option is excluded. So $a \in \min(y \setminus x)$. This shows that $x \cap y \subseteq \min(y \setminus x)$.

Now suppose $a \in \min(y \setminus x)$. Then $a \in y$ by transitivity of y. If $a \notin x$, then $a \in y \setminus x$, which would give $\min(y \setminus x) \,\underline{\in}\, a$, which is clearly impossible. So $a \in x$. This proves $\min(y \setminus x) \subseteq x \cap y$, and hence $\min(y \setminus x) = x \cap y$.

From Fact 1 we will prove the following.

Fact 2: *$x \,\underline{\in}\, y \Leftrightarrow x \subseteq y$ for any two $\mathfrak{U}$-ordinals x and y.*

To see why this holds, consider $\mathfrak{U}$-ordinals x and y such that $x \subseteq y$. Then $x = x \cap y$, and either $y \setminus x$ is empty, in which case $x = y$, or

$$x = x \cap y = \min(y \setminus x) \in y$$

by Fact 1. The other direction follows trivially from transitivity of y.

We now show that O is well-ordered under the relation $\underline{\in}$. First, we prove that O is totally ordered under $\underline{\in}$. Consider $x, y \in O$. If $x \neq y$, then either $x \setminus y$ or $y \setminus x$ is nonempty. Without loss of generality, suppose that $y \setminus x$ is nonempty. By Fact 1, $\mathsf{min}(y \setminus x) \subseteq x$. Then, by Fact 2, either $\mathsf{min}(y \setminus x) \in x$ or $\mathsf{min}(y \setminus x) = x$, and since $\mathsf{min}(y \setminus x) \in y \setminus x$, we can conclude that $x = \mathsf{min}(y \setminus x)$, and thus $x \in y$.

Now consider a nonempty $Y \subseteq O$, and any $y \in Y$. If $y \cap Y = \varnothing$, then for all $x \in Y$ we have $x \notin y$ and thus $y = \mathsf{min}\, Y$ by total ordering. If $y \cap Y \neq \varnothing$, then $\mathsf{min}(y \cap Y)$ exists, since $y \cap Y \subseteq y$. For all $x \in Y \setminus y$, it holds that $x \notin y$, and thus $y \subseteq x$ (by total ordering and Fact 2), which in turn implies $\mathsf{min}(y \cap Y) \in x$. We can conclude that $\mathsf{min}(y \cap Y) \in x$ for all $x \in Y$, and thus $\mathsf{min}\, Y = \mathsf{min}(y \cap Y)$.

We will now complete the proof of (c) by showing that $X^{<}$ is nonempty for all $X \in \mathsf{P}_{\mathfrak{U}}O$, and simultaneously prove that $\overset{+}{\bigvee} X = X \cup \bigcup X$ for all $X \in \mathsf{P}_{\mathfrak{U}}O$.

Consider $X \in \mathsf{P}_{\mathfrak{U}}O$. By (a) and Lemma 36, $X \in \mathfrak{U}$. Then $X \cup \bigcup X \in \mathfrak{U}$. To show that $X \cup \bigcup X$ is a $\mathfrak{U}$-ordinal, we need to prove that it is a transitive set, well-ordered under the relation $\underline{\in}$, and having no element z such that $z \in z$.

Since each element of X is transitive, $\bigcup\bigcup X \subseteq \bigcup X$ holds, which implies transitivity of $X \cup \bigcup X$:

$$\begin{aligned}\bigcup(X \cup \bigcup X) &= \bigcup X \cup \bigcup\bigcup X \\ &= \bigcup X \\ &\subseteq X \cup \bigcup X.\end{aligned}$$

Since O is transitive by (b), and $X \subseteq O$, each element of $\bigcup X$ is a $\mathfrak{U}$-ordinal, and thus each element of $X \cup \bigcup X$ is a $\mathfrak{U}$-ordinal. Then the fact that $X \cup \bigcup X$ is well-ordered under the relation $\underline{\in}$ follows from the fact that O is well-ordered under the same relation, as we have already proven. Also, since each element $z \in X \cup \bigcup X$ is a $\mathfrak{U}$-ordinal, we will never have $z \in z$.

Then, since $X \cup \bigcup X$ is transitive, well-ordered under $\underline{\in}$, and has no element z such that $z \in z$, we have that $X \cup \bigcup X$ is a $\mathfrak{U}$-ordinal.

Now, let us remark that for any $X \in \mathsf{P}_{\mathfrak{U}}O$,

$$\begin{aligned}\overset{+}{\bigvee} X &= \mathsf{min}\, X^{<} \\ &= \mathsf{min}\{y \in O \mid \forall_{x \in X}\, x \in y\} \\ &= \mathsf{min}\{y \in O \mid X \subseteq y\}.\end{aligned}$$

Notice that $X \cup \bigcup X \in \min X^{<}$, and consider any $y \in X^{<}$. Then $X \subseteq y$, and it is not hard to see that $X \cup \bigcup X \subseteq y$, since $\bigcup X \subseteq y$ by transitivity of y. Then, by Fact 2, $X \cup \bigcup X \in y$, and so

$$X \cup \bigcup X = \min X^{<} = \overset{+}{\bigvee} X.$$

Step 2: *We prove that if* (a–c) *hold, then O is the set of all $\mathfrak{U}$-ordinals.*

Let O be any set satisfying (a–c).

First we prove that every element of O is a $\mathfrak{U}$-ordinal, i.e. a transitive element of $\mathfrak{U}$ that is well-ordered by the relation $\underline{\in}$, and has no element z such that $z \in z$. Notice that the last fact holds by (c).

Let $x \in O$. By (c), the set O is ordered under the relation $\underline{\in}$. In this ordered set,

$$\{x\}^{>} = \{y \in O \mid y \in x\} = x \cap O.$$

By (b), $x \subseteq O$, and so $x = \{x\}^{>}$. By (a) and (c), $x \in \mathfrak{U}$. Furthermore, for any $y \in x$, the following holds:

$$y = \{y\}^{>} \subseteq \{x\}^{>} = x.$$

This shows that x is transitive.

Since O is well-ordered under $\underline{\in}$ (by (c) and Theorem 50) and $x \subseteq O$, we know x is also well-ordered under the same relation. Thus, every element of O is a $\mathfrak{U}$-ordinal.

Now we need to establish that every $\mathfrak{U}$-ordinal is in O. We already proved that the set $\mathcal{O}$ of all $\mathfrak{U}$-ordinals is an ordinal system, and since O is a set of $\mathfrak{U}$-ordinals, $O \subseteq \mathcal{O}$. The equality $O = \mathcal{O}$ then follows from Corollary 56. ■

2.5 Collecting systems and the corresponding universal property

In this section we introduce a notion of a 'collecting system'. The central idea is to take the property of an ordinal system that was established in Lemma 45,

$$\overset{+}{\bigvee} A = \bigvee A^{+},$$

and extend it to arbitrary increasing functions s:

$$\overset{s}{\bigvee} A = \bigvee sA.$$

This allows us to identify categories whose objects are collecting systems and whose morphisms are functions satisfying the identity

$$f(\overset{s}{\bigvee} A) = \overset{t}{\bigvee} fA.$$

We go on to prove that ordinal systems are initial objects in these categories. This universal property turns out to be one of the defining features of ordinal systems in Joyal and Moerdijk's book on algebraic set theory (see Chapter II: §2 of [7]) once we specialise their context to the context of this thesis.

Definition 60 ($\mathfrak{U}$-collecting system)**.** Given a universe $\mathfrak{U}$, a $\mathfrak{U}$-*collecting system* is a triple $X = (X, s, \leqslant)$ satisfying the following axioms.

- **(CA1)** $(X, \leqslant)$ is a join semi-lattice (see Definition 19) such that the join $\bigvee A$ is defined for each $A \in \mathsf{P}_{\mathfrak{U}}X$.
- **(CA2)** s is an increasing function $X \to X$ called the *incrementing* function of X.

We also define the *collecting function* $\overset{s}{\bigvee}\colon \mathsf{P}_{\mathfrak{U}}X \to X$ of X as follows: for all $A \in \mathsf{P}_{\mathfrak{U}}X$,

$$\overset{s}{\bigvee} A = \bigvee sA.$$

Remark 61. Note that for any ordinal system $(\mathcal{O}, \leqslant)$ relative to a universe $\mathfrak{U}$, there is a corresponding $\mathfrak{U}$-collecting system $(\mathcal{O}, \_^{+}, \leqslant)$, where $\_^{+}$ is the successor function of $\mathcal{O}$. By Lemma 45, its collecting function is the incremented join, $\overset{+}{\bigvee}$. We call such a $\mathfrak{U}$-collecting system an *ordinal* $\mathfrak{U}$-*collecting system*.

Remark 62. Consider $\mathfrak{U}$-collecting systems $(X, s, \leqslant_X)$, $(Y, t, \leqslant_Y)$ and $(Z, r, \leqslant_Z)$, and functions $f\colon X \to Y$ and $g\colon Y \to Z$ that satisfy, for all $A \in \mathsf{P}_{\mathfrak{U}}X$ and $B \in \mathsf{P}_{\mathfrak{U}}Y$,

$$f(\overset{s}{\bigvee} A) = \overset{t}{\bigvee} fA \quad \text{and} \quad g(\overset{t}{\bigvee} B) = \overset{r}{\bigvee} gB.$$

Then the composite $g \circ f$ satisfies, for all $A \in \mathsf{P}_{\mathfrak{U}}X$,

$$(g \circ f)(\overset{s}{\bigvee} A) = g(\overset{t}{\bigvee} fA) = \overset{r}{\bigvee}(g \circ f)A.$$

Furthermore, since f and g are set functions, composition is associative. It is easy to see that the identity function 1_X satisfy, for all $A \in \mathsf{P}_{\mathfrak{U}}X$,

$$1_X(\overset{s}{\bigvee} A) = \overset{s}{\bigvee} A = \overset{s}{\bigvee} 1_X A.$$

Definition 63 (category of $\mathfrak{U}$-collecting systems)**.** For a fixed universe $\mathfrak{U}$, the *category of $\mathfrak{U}$-collecting systems* is a category whose objects are the $\mathfrak{U}$-collecting systems, and whose morphisms are the functions $f\colon (X, \leqslant_X) \to (Y, \leqslant_Y)$ that satisfy, for all $A \in \mathsf{P}_{\mathfrak{U}}X$,

$$f(\bigvee^{s} A) = \bigvee^{t} fA.$$

Theorem 64 (the universal property of an ordinal $\mathfrak{U}$-collecting system)**.** *Let $(\mathcal{O}, \_^{+}, \leqslant)$ be an ordinal $\mathfrak{U}$-collecting system, and let $(X, s, \preccurlyeq)$ be any $\mathfrak{U}$-collecting system. Then there exists a unique function $f\colon \mathcal{O} \to X$ satisfying the following identity for every $A \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$:*

$$f(\bigvee^{+} A) = \bigvee^{s} fA.$$

Proof. Suppose at least one function $g\colon \mathcal{O} \to X$ exists such that for all $A \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$, we have $g(\bigvee^{+} A) = \bigvee^{s} gA$. Then g satisfies the following:

(a) For every ordinal x,

$$\begin{aligned} g(x^{+}) &= g(\bigvee^{+}\{x\}) \\ &= \bigvee^{s} g\{x\} \\ &= \bigvee\{s(g(x))\} \\ &= s(g(x)). \end{aligned}$$

(b) For every limit ordinal x,

$$\begin{aligned} g(x) &= g(\bigvee^{+}\{y \mid y < x\}) \\ &= \bigvee^{s} g\{y \mid y < x\}. \end{aligned}$$

Thus, if there is a function satisfying the identity, it must be the unique function f defined by the following transfinite recursion:

$$f(x^{+}) = s(f(x)); \qquad \text{(for every ordinal } x\text{)}$$

$$f(x) = \bigvee^{s} f\{y \mid y < x\}. \qquad \text{(for every limit ordinal } x\text{)}$$

First, we prove by transfinite induction that f is increasing.

Successor case: Suppose, for some ordinal x and all ordinals $z_1 \leqslant z_2 \leqslant x$, that $f(z_1) \preccurlyeq f(z_2)$.

Notice that for each ordinal $y < x$, we have $y^{+} \leqslant x$ by (L3), and thus, by induction assumption and since s is increasing,

$$f(y) \preccurlyeq f(y^{+}) = s(f(y)) \preccurlyeq s(f(x)) = f(x^{+}).$$

Thus, if x is a successor ordinal, $f(x) \preccurlyeq f(x^+)$.

If x is a limit ordinal, then $f(x^+)$ is an upper bound for $\{s(f(y)) \mid y < x\}$, as we just proved. Thus we must have

$$\begin{aligned} f(x) &= \overset{s}{\bigvee} f\{y \mid y < x\} \\ &= \bigvee\{s(f(y)) \mid y < x\} \\ &\preccurlyeq f(x^+). \end{aligned}$$

In both cases we get $f(x) \preccurlyeq f(x^+)$, and by induction assumption this gives us that $f(z_1) \preccurlyeq f(z_2)$ holds for all $z_1 \leqslant z_2 \leqslant x^+$, as required.

Limit case: Suppose, for some limit ordinal x and all ordinals $z_1 \leqslant z_2 < x$, that $f(z_1) \preccurlyeq f(z_2)$. Since

$$\begin{aligned} f(x) &= \overset{s}{\bigvee} f\{y \mid y < x\} \\ &= \bigvee sf\{y \mid y < x\} \\ &= \bigvee f\{y^+ \mid y < x\}, \end{aligned}$$

we have $f(y) \preccurlyeq f(y^+) \preccurlyeq f(x)$ for each $y < x$. By induction assumption, this means $f(z_1) \preccurlyeq f(z_2)$ for all $z_1 \leqslant z_2 \leqslant x$.

So f is increasing. Now we prove by transfinite induction on $\forall A$ that f satisfies the identity $f(\forall A) = \overset{s}{\bigvee} fA$ for all $A \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$.

Successor case: Suppose that for some $x \in \mathcal{O}$, and for all subsets B of $\mathcal{O}$ such that $\forall B = x$, it holds that $f(\forall B) = \overset{s}{\bigvee} fB$. Now consider any $A \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$ such that $\forall A = x^+$. Then $x = \max A$ (by Lemma 46), and thus $x^+ = \max A^+$ (by (L7)). Since f is increasing,

$$\begin{aligned} \overset{s}{\bigvee} fA &= \bigvee\{s(f(y)) \mid y \in A\} \\ &= \bigvee\{f(y^+) \mid y \in A\} \\ &= f(x^+) \\ &= f(\forall A). \end{aligned}$$

Limit case: Suppose that for some limit ordinal x, and for all subsets B of $\mathcal{O}$ such that $\forall B < x$, it holds that $f(\forall B) = \overset{s}{\bigvee} fB$. Now consider any subset A of $\mathcal{O}$ such that $\forall A = x$. Then $A \subseteq \{y \mid y < x\}$, and thus

$$sfA \subseteq sf\{y \mid y < x\}.$$

Then

$$
\begin{aligned}
\overset{s}{\bigvee} fA &= \bigvee sfA \\
&\preccurlyeq \bigvee sf\{y \mid y < x\} \\
&= \overset{s}{\bigvee} f\{y \mid y < x\} \\
&= f(x) \\
&= f(\overset{+}{\bigvee} A).
\end{aligned}
$$

Now, since $\overset{+}{\bigvee} A = x$, for each $z < x$ there exists $z' \in A$ such that $z \leqslant z'$, and, since s and f are increasing, $s(f(z)) \preccurlyeq s(f(z'))$. Then

$$
\begin{aligned}
f(\overset{+}{\bigvee} A) &= f(x) \\
&= \overset{s}{\bigvee} f\{y \mid y < x\} \\
&= \bigvee \{s(f(y)) \mid y < x\} \\
&\preccurlyeq \bigvee \{s(f(y)) \mid y \in A\} \\
&= \overset{s}{\bigvee} fA.
\end{aligned}
$$

Thus, since $\overset{s}{\bigvee} fA \preccurlyeq f(\overset{+}{\bigvee} A) \preccurlyeq \overset{s}{\bigvee} fA$, we have $f(\overset{+}{\bigvee} A) = \overset{s}{\bigvee} fA$.

This proves that $f(\overset{+}{\bigvee} A) = \overset{s}{\bigvee} fA$ holds for all $A \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$, and we have already remarked that f must be the unique function with this property. ∎

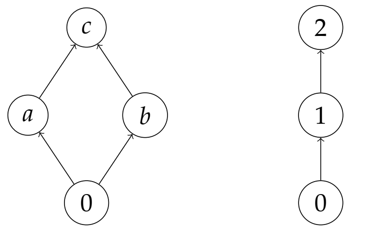
Remark 65. Consider $\mathfrak{U}$-collecting systems $(X, s, \leqslant_X)$ and $(Y, t, \leqslant_Y)$ and a function $f\colon X \to Y$. While, in general,

$$
f(\overset{s}{\bigvee} A) = f(\bigvee sA) = \bigvee tfA = \overset{t}{\bigvee} fA
$$

holds for all $A \in \mathsf{P}_{\mathfrak{U}}X$ when f preserves both joins and successors, the converse is not true. Successors will always be preserved by morphisms of $\mathfrak{U}$-collecting systems, since

$$
\begin{aligned}
f(s(x)) &= f(\bigvee s\{x\}) \\
&= f(\overset{s}{\bigvee}\{x\}) \\
&= \overset{t}{\bigvee} f\{x\} \\
&= \bigvee tf\{x\} \\
&= t(f(x)).
\end{aligned}
$$

However, joins might not be preserved – we will give a simple example. Consider the following join semi-lattices:



$(X, \leqslant_X)$ $(Y, \leqslant_Y)$

We can define incrementing functions $s\colon X \to X$ and $t\colon Y \to Y$ by

$$s(x) = \begin{cases} a & \text{if } x = 0, \\ c & \text{otherwise,} \end{cases} \qquad t(y) = \begin{cases} 1 & \text{if } y = 0, \\ 2 & \text{otherwise.} \end{cases}$$

Since $(X, s, \leqslant_X)$ and $(Y, t, \leqslant_Y)$ are complete lattices and s and t are increasing, they are $\mathfrak{U}$-collecting systems for any universe $\mathfrak{U}$. We define a function $f\colon X \to Y$ such that

$$f(0) = 0; \qquad f(a) = f(b) = 1; \qquad f(c) = 2.$$

It is not difficult to see that f is a morphism of $\mathfrak{U}$-collecting systems. Taking $\{a, b\}$ as an example,

$$\begin{gathered} f\big(\overset{s}{\bigvee}\{a,b\}\big) = f\big(\bigvee\{c\}\big) = f(c) = 2, \\ \overset{t}{\bigvee} f\{a,b\} = \overset{t}{\bigvee}\{1\} = \bigvee\{2\} = 2 = f\big(\overset{s}{\bigvee}\{a,b\}\big), \end{gathered}$$

but

$$\begin{gathered} f\big(\bigvee\{a,b\}\big) = f(c) = 2, \\ \bigvee f\{a,b\} = \bigvee\{1\} = 1 \neq f\big(\bigvee\{a,b\}\big), \end{gathered}$$

i.e. f preserves incremented joins but not joins.

Chapter 3

Counting systems

In this chapter we develop an alternative, entirely new approach to the ordinal number system, which mimics Dedekind's approach to the natural number system recalled in Section 1.4. The structural base for this approach is given, in the first stage, by the notion of a 'limit-successor system' in Section 3.1, and, in the second stage, by a 'counting system' in Section 3.2. The characterisation of ordinal systems via Dedekind-style axioms is given by Theorem 78 in Section 3.1. This approach to ordinal systems leads to a universal property of its own, which is formulated and proved in Section 3.3.

Note that the 'successor functions' in this chapter are generally different from the successor function in Chapter 2, but turn out to be equivalent when Axioms (C1–5) are satisfied.

3.1 Limit-successor systems

Definition 66 (limit-successor system)**.** A *limit-successor system* is a triple $X = (X, L, s)$, where X is a set, L is a partial function $L\colon \mathsf{P}X \to X$ called the *limit function*, and s is a function $s\colon X \to X$ called the *successor function*.

Definition 67 (successor-closed subset)**.** A *successor-closed* subset of a limit-successor system X is a subset I of X such that $sI \subseteq I$.

Notation 68. For a subset of a limit-successor system X, we write $s^{-1}I$ to denote

$$s^{-1}I = \{x \in X \mid s(x) \in I\},$$

and $L^{-1}I$ to denote

$$L^{-1}I = \{A \in \operatorname{dom} L \mid L(A) \in I\}.$$

Recall that a topology τ on X is called an *Alexandrov topology*, and (X, τ) an *Alexandrov-discrete space*, if $\bigcap A \in \tau$ for each $A \subseteq \tau$ (originally due to Pavel Alexandrov in 1936 [26]; also see e.g. [27, 28]). Thus, $\tau \subseteq \mathrm{P}X$ must satisfy the following three conditions:

(a) $X, \varnothing \in \tau$;

(b) $\bigcup A \in \tau$ for each $A \subseteq \tau$;

(c) $\bigcap A \in \tau$ for each $A \subseteq \tau$.

Note that, because of the way we defined the intersection of a set in this thesis, $\bigcap \varnothing = \varnothing$ (see Section 1.2), and thus we do not need to specify that A must be nonempty in the case of (c), as many authors do.

Recall further, that in an Alexandrov-discrete space (X, τ), the relation $\leqslant$ such that for all $x, y \in X$,

$$x \leqslant y \quad \Leftrightarrow \quad x \in \overline{\{y\}},$$

is a preorder (i.e. reflexive and transitive – see Definition 10), called the *specialisation preorder* on X.

Lemma 69. *Let (X, L, s) be a limit-successor system and let*

$$\tau' = \{I \subseteq X \mid [s^{-1}I \subseteq I] \wedge [\bigcup L^{-1}I \subseteq I]\}.$$

Then τ' is an Alexandrov topology on X.

Proof. We check the three conditions for τ' to be an Alexandrov topology.

a. Clearly, $s^{-1}X \subseteq X$ and $\bigcup L^{-1}X \subseteq X$ hold, since s and L only deal with elements of X. Since $s(x) \in \varnothing$ and $L(A) \in \varnothing$ are never true, it also holds that $s^{-1}\varnothing \subseteq \varnothing$ and $\bigcup L^{-1}\varnothing \subseteq \varnothing$. Thus, $X, \varnothing \in \tau'$.

b. Let $A \subseteq \tau'$, and consider any $x \in s^{-1}\bigcup A$. Then, by definition of s^{-1}, we have $s(x) \in \bigcup A$, and thus for some $I \in A$, we have $s(x) \in I$ and thus $x \in s^{-1}I$. Since $I \in \tau'$, this gives us

$$x \in s^{-1}I \subseteq I \subseteq \bigcup A.$$

Now consider $x \in \bigcup L^{-1}\bigcup A$. Then $x \in S$ for some $S \in L^{-1}\bigcup A$. Then $L(S) \in \bigcup A$ by definition of L^{-1}, and thus $L(S) \in I$ for some $I \in A$. Then $S \in L^{-1}I$ by definition of L^{-1}, and thus $S \subseteq \bigcup L^{-1}I$. Since $I \in \tau'$, this means that

$$x \in S \subseteq \bigcup L^{-1}I \subseteq I \subseteq \bigcup A.$$

Thus, $s^{-1}\bigcup A \subseteq \bigcup A$ and $\bigcup L^{-1}\bigcup A \subseteq \bigcup A$ both hold, giving us $\bigcup A \in \tau'$.

c. Let $A \subseteq \tau'$, and consider any $x \in s^{-1} \bigcap A$. By definition of s^{-1}, we have $s(x) \in \bigcap A$, i.e. for each $I \in A$, we have $s(x) \in I$. Then for an arbitrary $I \in A$, we have $x \in s^{-1}I$ by definition of s^{-1}. Since $I \in \tau'$, this means

$$x \in s^{-1}I \subseteq I.$$

So $x \in I$ holds for each $I \in A$, which gives us $x \in \bigcap A$.

Now consider any $x \in \bigcup L^{-1} \bigcap A$. Then $x \in S$ for some $S \in L^{-1} \bigcap A$. Then $L(S) \in \bigcap A$ by definition of L^{-1}, i.e. $L(S) \in I$ holds for each $I \in A$. Then, for an arbitrary $I \in A$, we have $S \in L^{-1}I$ by definition of L^{-1}, and thus, since $I \in \tau'$,

$$x \in S \subseteq \bigcup L^{-1}I \subseteq I.$$

So $x \in I$ holds for each $I \in A$, which gives us $x \in \bigcap A$.

Thus, $s^{-1} \bigcap A \subseteq \bigcap A$ and $\bigcup L^{-1} \bigcap A \subseteq \bigcap A$ both hold, giving us $\bigcap A \in \tau'$.

This proves that τ' is indeed an Alexandrov topology on X. ■

Note, for the following definition, that if τ' is any Alexandrov topology on X, then so is

$$\tau = \{X \setminus A \mid A \in \tau'\}.$$

Then a set is open in τ' if and only if it is closed in τ.

Definition 70 (limit-successor system topology)**.** Let $X = (X, L, s)$ be a limit-successor system. We define the *limit-successor system topology* τ on X as the topology whose closed sets are the subsets I of X that satisfy

$$s^{-1}I \subseteq I \quad \text{and} \quad \bigcup L^{-1}I \subseteq I.$$

Notation 71. We abbreviate the operator s^{-1} composed with itself m times as s^{-m}, with the $m = 0$ case giving the identity operator. We write $s^{-\infty}$ for the operator defined by

$$s^{-\infty}I = \bigcup\{s^{-m}I \mid m \in \mathbb{N}\},$$

and $\bigcup L^{-1}$ for the operator $I \mapsto \bigcup L^{-1}I$.

Theorem 72. *The closure of a subset I of a limit-successor system X can be computed as*

$$\bar{I} = \bigcup\{s^{-\infty}[\bigcup L^{-1}s^{-\infty}]^k I \mid k \in \mathbb{N}\}.$$

Proof. Consider

$$J = \bigcup\{s^{-\infty}[\bigcup L^{-1}s^{-\infty}]^k I \mid k \in \mathbb{N}\}.$$

We want to prove that J is closed, and that $J \subseteq K$ for every closed K such that $I \subseteq K$.

Step 1: *We prove that J is closed.*

First we prove that $s^{-1}J \subseteq J$. Consider any $x \in s^{-1}J$. Then

$$\begin{aligned} x \in s^{-1}J &= s^{-1}\bigcup\{s^{-\infty}[\bigcup L^{-1}s^{-\infty}]^k I \mid k \in \mathbb{N}\} \\ &= \{y \in X \mid s(y) \in \bigcup\{s^{-\infty}[\bigcup L^{-1}s^{-\infty}]^k I \mid k \in \mathbb{N}\}\}. \end{aligned}$$

Then for some $n \in \mathbb{N}$, using the fact that $s^{-\infty}[\bigcup L^{-1}s^{-\infty}]^n I$ is closed under s^{-1},

$$\begin{aligned} x \in s^{-1}s^{-\infty}[\bigcup L^{-1}s^{-\infty}]^n I \subseteq s^{-\infty}[\bigcup L^{-1}s^{-\infty}]^n I \\ \subseteq \bigcup\{s^{-\infty}[\bigcup L^{-1}s^{-\infty}]^k I \mid k \in \mathbb{N}\} \\ = J. \end{aligned}$$

Thus, $s^{-1}J \subseteq J$. Now consider $x \in X$ such that

$$x \in \bigcup L^{-1}J = \bigcup L^{-1}\bigcup\{s^{-\infty}[\bigcup L^{-1}s^{-\infty}]^k I \mid k \in \mathbb{N}\}.$$

Then for some $n \in \mathbb{N}$,

$$\begin{aligned} x \in \bigcup L^{-1}s^{-\infty}[\bigcup L^{-1}s^{-\infty}]^n I &= [\bigcup L^{-1}s^{-\infty}]^{n+1} I \\ &\subseteq s^{-\infty}[\bigcup L^{-1}s^{-\infty}]^{n+1} I \\ &\subseteq \bigcup\{s^{-\infty}[\bigcup L^{-1}s^{-\infty}]^k I \mid k \in \mathbb{N}\} \\ &= J. \end{aligned}$$

We can conclude that $\bigcup L^{-1}J \subseteq J$, and thus J is closed.

Step 2: *We prove that $J \subseteq H$ for any closed H such that $I \subseteq H$.*

Consider any closed H such that $I \subseteq H$. Then for each $K \subseteq H$,

$$s^{-1}K \subseteq s^{-1}H \subseteq H.$$

Then by induction, for each $n \in \mathbb{N}$ we have

$$s^{-n}K \subseteq H,$$

and thus

$$s^{-\infty}K = \bigcup\{s^{-m}K \mid m \in \mathbb{N}\} \subseteq H.$$

Similarly, for each $K \subseteq H$, we have

$$\bigcup L^{-1}K \subseteq \bigcup L^{-1}H \subseteq H.$$

Combining these facts, we have for each $K \subseteq H$,

$$\bigcup L^{-1}s^{-\infty}K \subseteq H.$$

Then, again by induction, we get that the following holds for all $n \in \mathbb{N}$:

$$s^{-\infty}[\bigcup L^{-1}s^{-\infty}]^n I \subseteq H.$$

This gives us

$$J = \bigcup\{s^{-\infty}[\bigcup L^{-1}s^{-\infty}]^k I \mid k \in \mathbb{N}\} \subseteq H.$$

Then, since J is closed by Step 1, we can conclude that J is the closure of I. ■

This theorem means that the specialisation preorder 'breaks up' into two relations $\leqslant_s$ and $\leqslant_L$, each determined by s and L alone, as explained in what follows. These relations are defined by:

$$\begin{array}{lcll} x \leqslant_s y & \Leftrightarrow & \exists_{m\in\mathbb{N}}[s^m(x) = y] & (\Leftrightarrow x \in s^{-\infty}\{y\}), \\ x \leqslant_L y & \Leftrightarrow & \exists_I[[x \in I] \wedge [L(I) = y]] & (\Leftrightarrow x \in \bigcup L^{-1}\{y\}). \end{array}$$

We then have $x \leqslant y$ if and only if

$$x = z_0 \leqslant_s z_1 \leqslant_L z_2 \leqslant_s z_3 \leqslant_L \ldots z_{2k} \leqslant_s y$$

for some $z_0, \ldots, z_{2k} \in X$, where k can be any natural number. Note that $\leqslant_s$ is both reflexive and transitive, although the same cannot be claimed for $\leqslant_L$.

3.2 Counting systems

Definition 73 ($\mathfrak{U}$-counting system)**.** Given a universe $\mathfrak{U}$, a $\mathfrak{U}$-*counting system* is a limit-successor system (X, L, s) satisfying the following conditions.

(C1) The domain of L is the set of all successor-closed subsets $I \in \mathsf{P}_{\mathfrak{U}}X$.

(C2) If I and J belong to the domain of L and $\overline{I} = \overline{J}$, then $L(I) = L(J)$.

The structure above is the one that will be used for formulating the universal property of an ordinal system in the next section.

Lemma 74. *Let $\mathfrak{U}$ be a universe and let (X, L, s) be a $\mathfrak{U}$-counting system. Then $\leqslant_L$ is transitive and furthermore,*

$$x \leqslant_s y \leqslant_L z \quad \Rightarrow \quad x \leqslant_L z$$

for all $x, y, z \in X$.

Proof. To prove transitivity of $\leqslant_L$, suppose $x \leqslant_L y$ and $y \leqslant_L z$. Then $x \in I$, $L(I) = y$, $y \in J$ and $L(J) = z$ for some successor-closed $I, J \in \mathsf{P}_{\mathfrak{U}}X$. The union $I \cup J \in \mathsf{P}_{\mathfrak{U}}X$ (Lemma 33) is also successor-closed, and hence it belongs to the domain of L by (C1). Since $L(I) \in J$, we get that $I \subseteq \overline{J}$. This implies that $\overline{I \cup J} = \overline{J}$. By (C2), $L(I \cup J) = L(J)$. Having $x \in I \cup J$ and $L(I \cup J) = z$ means that $x \leqslant_L z$. This completes the proof of transitivity.

To prove the second property, suppose $x \leqslant_s y \leqslant_L z$. Then $s^m(x) = y$, and $y \in J$ where $L(J) = z$, for some $m \in \mathbb{N}$ and some successor-closed $J \in \mathsf{P}_{\mathfrak{U}}X$. Here we also expand J, this time adding to it all elements of the form $s^k(x)$, where $k \in \{0, \ldots, m-1\}$. The resulting set,

$$K = \{x, s(x), \ldots, s^{m-1}(x)\} \cup J,$$

is clearly successor-closed and belongs to $\mathsf{P}_{\mathfrak{U}}X$ (Lemma 31 and Lemma 33). Then, since $\overline{K} = \overline{J}$, we get that $L(K) = L(J)$ by (C2). This implies $x \leqslant_L z$. ■

This lemma gives that in a $\mathfrak{U}$-counting system (X, L, s), for any $x, z \in X$ we have

$$x \leqslant z \quad \Leftrightarrow \quad [x \leqslant_s z] \vee \exists_y [x \leqslant_L y \leqslant_s z].$$

Corollary 75. *The successor function of a $\mathfrak{U}$-counting system (X, L, s) is increasing under the specialisation preorder $\leqslant$ of X.*

Proof. Consider $x, y \in X$ such that $x \leqslant y$. By Lemma 74, we need to consider two cases.

Case 1: $x \leqslant_s y$. Then for some $k \in \mathbb{N}$ we have $s^k(x) = y$, and thus,

$$s^k(s(x)) = s^{k+1}(x) = s(s^k(x)) = s(y).$$

This gives us $s(x) \leqslant_s s(y)$.

Case 2: $x \leqslant_L z \leqslant_s y$ for some $z \in X$. Then, using Case 1,

$$s(x) \leqslant_L z \leqslant_s y \leqslant_s s(y),$$

since L is defined only for successor-closed sets.

In both cases, $s(x) \leqslant s(y)$, and thus s is increasing. ∎

Corollary 76. *If $x < y$ then $s(x) \leqslant y$, for all $x, y \in X$.*

Proof. Suppose $x < y$. There are two cases:

Case 1: $x \leqslant_s y$. Then clearly $s(x) \leqslant y$, since $x \neq y$.

Case 2: $x \leqslant_L y' \leqslant_s y$ for some y'. Then $x \in I$ and $L(I) = y'$ for some successor-closed I, by (C1). So $s(x) \in I$, and thus $s(x) \leqslant_L y'$. With $y' \leqslant_s y$, this gives $s(x) \leqslant y$.

So in both cases we get $s(x) \leqslant y$, as required. ∎

Lemma 77. *In a $\mathfrak{U}$-counting system (X, L, s), the closure of a subset $I \subseteq X$ is given by*

$$\bar{I} = s^{-\infty} I \cup \bigcup L^{-1} s^{-\infty} I.$$

Proof. Consider

$$J = s^{-\infty} I \cup \bigcup L^{-1} s^{-\infty} I.$$

Since J is certainly a subset of

$$\bar{I} = \bigcup \{ s^{-\infty} [\bigcup L^{-1} s^{-\infty}]^k I \mid k \in \mathbb{N} \},$$

we only need to prove that J is closed. This will be the case if J is down-closed under $\leqslant$.

Consider any $y \in J$, and any $x \in X$ such that $x \leqslant y$. We consider two cases.

Case 1: $x \leqslant_s y$. Then

$$x \in s^{-\infty}\{y\} \subseteq s^{-\infty} I \subseteq J.$$

Case 2: $x \leqslant_L z \leqslant_s y$ for some $z \in X$. Then $z \in s^{-\infty} I$, as noted in Case 1. Thus, since $x \leqslant_L z$, we have

$$x \in \bigcup L^{-1}\{z\} \subseteq \bigcup L^{-1} s^{-\infty} I \subseteq J.$$

Either way, $x \in J$. So, since J is closed and $I \subseteq J \subseteq \bar{I}$, we have $J = \bar{I}$. ∎

Theorem 78. *The specialisation preorder $\leqslant$ of a $\mathfrak{U}$-counting system (X, L, s) makes $(X, \leqslant)$ an ordinal system relative to $\mathfrak{U}$, provided the following conditions hold.*

(C3) $s^{-1}\{L(I)\} = \varnothing = I \cap \{s(L(I))\}$ *for all* I *such that* $L(I)$ *is defined.*

(C4) s *is injective and* L *has the property that if* $L(I) = L(J)$ *then* $\overline{I} = \overline{J}$.

(C5) $J = X$ *for any successor-closed set* J *having the property that* $I \subseteq J \Rightarrow L(I) \in J$ *every time* $L(I)$ *is defined.*

When these conditions hold, s *is the successor function of the ordinal system and* $L(I) = \bigvee I = \curlyvee I$ *whenever* $L(I)$ *is defined; moreover, the limit ordinals are exactly the elements of* X *of the form* $L(I)$*. Furthermore, the closure of* $I \in \mathsf{P}_{\mathfrak{U}}X$ *is given by* $\overline{I} = \{\curlyvee I\}^{>}$*. Finally, any ordinal system relative to* $\mathfrak{U}$ *arises this way from a (unique)* $\mathfrak{U}$*-counting system satisfying* (C3–5).

Before proving the theorem, let us illustrate axioms (C1–5) in the case where $\mathfrak{U}$ is the universe of hereditarily finite sets. Then L is only defined for finite successor-closed sets. Consider such a set I. Since s is injective and I is finite, each $x \in I$ must have the property that $x = s^k(x)$ for some natural number $k > 0$. By (C3), such x cannot lie in the image of L.

Now consider

$$J = \{x \in X \mid \forall_{k>0}[s^k(x) \neq x]\}.$$

It is easy to see that J is successor-closed. Since no x such that $s^k(x) = x$ for some $k \in \mathbb{N}$ can lie in the image of L, we have $L(I) \in J$ for every successor-closed $I \subseteq J$, and thus $J = X$ by (C5). It follows that L can only be defined for the empty set, and thus serves to select a distinguished element of X. The triple (X, L, s) then becomes a triple $(X, 0, s)$, where $0 = L(\varnothing)$ is the unique element in the image of L. The axioms (C1–2) then trivially hold, while (C3–5) take the form of the axioms of Dedekind for a natural number system:

- The first equality in (C3) states that 0 does not belong to the image of s, while the second equality holds trivially.
- (C4) just states that s is injective.
- (C5) becomes the usual principle of mathematical induction.

Proof of Theorem 78. Suppose conditions (C1–5) hold.

Step 1: *As a first step, we prove that the specialisation preorder is antisymmetric, i.e. that it is a partial order.*

For this, we first show that $\leqslant_L$ is 'antireflexive': it is impossible to have $x \leqslant_L x$. Indeed, suppose $x \in I$ and $L(I) = x$. Since I is successor-closed by (C1), $s(x) \in I$. But then $s(x) \in I \cap \{s(L(I))\}$, which is impossible by (C3).

Next, we show antisymmetry of $\leqslant_s$. Suppose $x \leqslant_s z \leqslant_s x$ and $x \neq z$. Then we get that $s^k(x) = x$ for $k > 1$. We will now show that this is not possible. In fact, we establish a slightly stronger property, which will be useful later on as well:

Property 1: *$s^k(x) \neq x$ for all $x \in X$ and all natural numbers $k > 0$.*

Actually, we have already established this property in the remark after the theorem. Here is a more detailed argument. Consider the set

$$J = \{x \in X \mid \forall_{k>0}[s^k(x) \neq x]\}.$$

We will use (C5) to show that $J = X$.

First, we show that J is successor-closed. Let $y \in J$, and suppose $s^k(s(y)) = s(y)$ for some $k > 0$. Then, by injectivity of s (which is required in (C4)), $s^{k-1}(s(y)) = y$, which is impossible. So $s(y) \in J$, showing that J is successor-closed.

Now let $I \in \mathsf{P}_{\mathfrak{U}}X$ be a successor-closed subset of J (by (C1), I is such if and only if $L(I)$ is defined). Then $L(I) \neq s^k(L(I))$ for all $k > 0$ by the first equality in (C3). Thus $L(I) \in J$, and we can apply (C5) to get $J = X$, as desired. Antisymmetry of $\leqslant_s$ has thus been established.

We are now ready to prove the antisymmetry of $\leqslant$. Suppose $x \leqslant z$ and $z \leqslant x$. There are four cases to consider:

Case 1: $x \leqslant_s z \leqslant_s x$. Then $x = z$ by antisymmetry of $\leqslant_s$.

Case 2: $x \leqslant_L y \leqslant_s z \leqslant_s x$ for some y. Then

$$\begin{aligned} & y \leqslant_s z \leqslant_s x \leqslant_L y \\ \Rightarrow\ & y \leqslant_s x \leqslant_L y && \text{(by transitivity of } \leqslant_s) \\ \Rightarrow\ & y \leqslant_L y, && \text{(by Lemma 74)} \end{aligned}$$

which we have shown to be impossible.

Case 3: $x \leqslant_s z \leqslant_L y \leqslant_s x$ for some y. Similarly, in this case we get the impossible

$$\begin{aligned} & y \leqslant_s x \leqslant_s z \leqslant_L y \\ \Rightarrow\ & y \leqslant_s z \leqslant_L y && \text{(by transitivity of } \leqslant_s) \\ \Rightarrow\ & y \leqslant_L y. && \text{(by Lemma 74)} \end{aligned}$$

Case 4: $x \leqslant_L y \leqslant_s z \leqslant_L y' \leqslant_s x$ for some y, y'. In this case, too, we get the impossible

$$\begin{aligned} & y \leqslant_s z \leqslant_L y' \leqslant_s x \leqslant_L y \\ \Rightarrow \quad & y \leqslant_L y' \leqslant_L y && \text{(by Lemma 74)} \\ \Rightarrow \quad & y \leqslant_L y. && \text{(by transitivity of } \leqslant_L \text{ from Lemma 74)} \end{aligned}$$

We have thus shown that the specialisation preorder is antisymmetric. We will now establish the following property, which will be useful later on.

Property 2: *If $x < s(y)$ then $x \leqslant y$, for all $x, y \in X$.*

Suppose $x < s(y)$. Again, we have two cases:

Case 1: $x \leqslant_s s(y)$. This, together with $x \neq s(y)$, gives $x \leqslant y$ by injectivity of s from (C4).

Case 2: $x \leqslant_L x' \leqslant_s s(y)$ for some $x' \in X$. Since $x' \neq s(y)$ by the first equality in (C3), by injectivity of s (from (C4)), we must have $x' \leqslant_s y$. This gives us $x \leqslant y$.

We get $x \leqslant y$ in both cases.

Step 2: *We prove that the specialisation preorder is a total order.*

We will prove this by simultaneously establishing the following:

Property 3: *Let $y \in X$, and let $I \subseteq X$. If $x < y$ for all $x \in I$ and $L(I)$ is defined, then $L(I) \leqslant y$.*

Let

$$J = \{x \in X \mid \forall_{y \in X}[y \leqslant x \vee x \leqslant y]\}.$$

We will use (C5) to show that $J = X$. For this, we first prove that J is successor-closed. Let $x \in J$, and consider any $y \in X$. Since $x \leqslant s(x)$, if $y \leqslant x$, then $y \leqslant s(x)$. If $x < y$, then by Corollary 76, $s(x) \leqslant y$. This proves that J is successor-closed.

Now consider $L(I)$, where $I \subseteq J$. To prove that $L(I) \in J$, we proceed as follows. Let $x \in X$. If $x \leqslant y$ for at least one $y \in I$, then $x \leqslant L(I)$, since $y \leqslant_L L(I)$. Thus, it suffices to prove that the set

$$K_I = \{x \in X \mid [\forall_{y \in I}\, y < x] \Rightarrow [L(I) \leqslant x]\}$$

is the entire $K_I = X$. This we prove using (C5).

First, we show that K_I is successor-closed. Suppose $x \in K_I$. If $y < s(x)$ for all $y \in I$, then by Property 2, $y \leqslant x$ for all $y \in I$. If $x \in I$, then, since I is successor-closed by (C1), we will have $s(x) \in I$, which will violate the assumption that $y < s(x)$ for all $y \in I$. So we get that $y < x$ for all $y \in I$. Then $L(I) \leqslant x$, since $x \in K_I$. This implies $L(I) \leqslant s(x)$, thus proving that K_I is successor-closed.

Now let $H \subseteq K_I$ be such that $L(H)$ is defined. Suppose $y < L(H)$ for all $y \in I$. From the first equality in (C3), we get that $y \leqslant_L L(H)$ for each $y \in I$. So for each $y \in I$, there exists G_y such that $y \in G_y$ and $L(G_y) = L(H)$. This implies that $\overline{G_y} = \overline{H}$ for each $y \in I$ (by (C4)), and so $I \subseteq \overline{H}$. Since $I \subseteq J$, each element of H is comparable with each element of I. We consider two cases.

Case 1: for every $h \in H$, there exists $y \in I$ such that $h \leqslant y$. Then $H \subseteq \overline{I}$. This would then give $\overline{I} = \overline{H}$, and so by (C2), $L(I) = L(H)$, showing that $L(I) \leqslant L(H)$, as desired.

Case 2: there exists $h \in H$ such that $y < h$ for every $y \in I$. Then, since $H \subseteq K_I$, we have $L(I) \leqslant h$. This, together with $h \leqslant_L L(H)$, will give $L(I) \leqslant L(H)$.

In both cases, $L(I) \leqslant L(H)$. We have thus shown that K_I has the required properties for us to apply (C5) to conclude that $K_I = X$. This, then, shows that $L(I) \in J$, and so J has the required properties to conclude that $J = X$. The proof that the specialisation preorder is a total order is now complete. At the same time, since $J = X$, and for each $I \subseteq J$ such that $L(I)$ is defined, $K_I = X$, we have also established Property 3.

Step 3: *We now show that $L(I) = \bigvee I = \curlyvee I$ whenever $L(I)$ is defined, and that $s(x) = x^+$ for all $x \in X$.*

Property 1 and Property 3 show that $L(I)$ is the join of I for any I such that $L(I)$ is defined. Indeed, if $x \leqslant y$ for all $x \in I$, then for each $x \in I$, we also have $s(x) \leqslant y$. Since $x < s(x)$, as clearly $x \leqslant s(x)$ and by Property 1, $x \neq s(x)$, we get that $x < y$ for all $x \in I$. Then, by Property 3, $L(I) \leqslant y$, thus showing that $L(I)$ is a join of I.

Furthermore, when $L(I)$ is defined, I is successor-closed and so it cannot have a largest element, by Property 1. Then the join $L(I)$ of I must also be the incremented join of I (Lemma 46).

Finally, the property $x < s(x)$, together with Corollary 76, implies that $s(x) = x^+$ for each $x \in X$. Thus, once we prove that X is an ordinal system under the specialisation preorder, we have that s is its successor function and L is given by the join.

Step 4: *We prove that X is an ordinal system under the specialisation preorder, where limit ordinals are exactly the elements of the form $L(I)$.*

Consider the set

$$J = \{x \in X \mid \{x\}^{>} \in \mathsf{P}_{\mathfrak{U}} X\}.$$

If $x \in J$, then by Property 2, $\{s(x)\}^{>} = \{x\}^{>} \cup \{x\} \in \mathsf{P}_{\mathfrak{U}}X$ (Lemma 31 and Lemma 33), and so $s(x) \in J$. For any $x \in J$, we therefore have

$$\begin{aligned} \overline{\{x\}} &= \{y \in X \mid y \leqslant x\} \\ &= \{y \in X \mid y < s(x)\} && \text{(by Property 2)} \\ &= \{s(x)\}^{>} \in \mathsf{P}_{\mathfrak{U}}X. \end{aligned}$$

Suppose $I \subseteq J$ is such that $L(I)$ is defined, and define a function $f\colon I \to \mathsf{P}_{\mathfrak{U}}X$ by $f(x) = \overline{\{x\}}$. Then, by Lemma 33,

$$\bigcup fI = \bigcup\{\overline{\{x\}} \mid x \in I\} \in \mathsf{P}_{\mathfrak{U}}X.$$

Notice that $\overline{I} \subseteq \{\bigvee I\}^{>}$, since $\overline{I}$ is the down-closure of I under the specialisation preorder.

Then, note that $\overline{\bigcup I} = \bigcup\{\overline{x} \mid x \in I\}$ holds for all $I \subseteq X$, as is true in any Alexandrov topology, and thus

$$\begin{aligned} \{L(I)\}^{>} &= \{\bigvee I\}^{>} && \text{(by Step 3)} \\ &= \overline{I} && \text{(since } X \text{ is totally ordered)} \\ &= \overline{\bigcup\{\{x\} \mid x \in I\}} \\ &= \bigcup\{\overline{\{x\}} \mid x \in I\} \in \mathsf{P}_{\mathfrak{U}}X, \end{aligned}$$

and so $L(I) \in J$. By (C5), $J = X$, and thus (O1′) holds.

To prove that X is an ordinal system under the specialisation preorder, it remains to show that it satisfies (O2), i.e. that for any $Y \in \mathsf{P}_{\mathfrak{U}}X$, the incremented join of Y exists in X. If Y has a largest element, then the successor of that element is the incremented join of Y (Lemma 46). If $\mathfrak{U}$ does not contain an infinite set, then Y is finite, and so it has a largest element.

Now consider $Y \in \mathsf{P}_{\mathfrak{U}}X$ that has no largest element, with $\mathfrak{U}$ containing an infinite set. We define:

$$s^{\infty}Y = \{s^n(x) \mid [x \in Y] \wedge [n \in \mathbb{N}]\}.$$

This is, of course, the closure of Y under s. Consider a function $f\colon Y \times \mathbb{N} \to X$ defined by $f(x,n) = s^n(x)$. Then, by Lemma 34 and Lemma 35,

$$\begin{aligned} s^{\infty}Y &= \{s^n(x) \mid [x \in Y] \wedge [n \in \mathbb{N}]\} \\ &= \{f(x,n) \mid (x,n) \in Y \times \mathbb{N}\} \\ &= f(Y \times \mathbb{N}) \in \mathsf{P}_{\mathfrak{U}}X. \end{aligned}$$

Thus $L(s^\infty Y)$ is defined, by (C1). Since $Y \subseteq s^\infty Y$, it holds that $y < L(s^\infty Y)$ for all $y \in Y$. Let $y \in Y$. We prove by induction on n that for each $n \in \mathbb{N}$, we have $s^n(y) < z$ for some $z \in Y$.

Base case: If $n = 0$, this follows from the fact that Y does not have a largest element.

Induction step: Suppose $s^n(y) < z$ for some $z \in Y$. Then $s^{n+1}(y) = s(s^n(y)) \leqslant z$. Since z cannot be the largest element of Y, we must have $z < z'$ for some $z' \in Y$. Then $s^{n+1}(y) < z'$.

What we have shown implies that the incremented join $L(s^\infty Y)$ of $s^\infty Y$ is also the incremented join of Y.

We have thus proven that the specialisation preorder of a $\mathfrak{U}$-counting system satisfying (C3–5) makes it an ordinal system relative to $\mathfrak{U}$, with s as its successor function, and L given equivalently by join and by incremented join. This also establishes that, if an ordinal system relative to $\mathfrak{U}$ arises this way from a $\mathfrak{U}$-counting system satisfying (C3–5), then this $\mathfrak{U}$-counting system is unique.

We now prove the existence of such a $\mathfrak{U}$-counting system. Actually, before doing that, note that by (C3), no element of X of the form $L(I)$ can be a successor ordinal, and so it must be a limit ordinal. Conversely, for a limit ordinal x we have $x = L(\{x\}^>)$ (Lemma 52). This shows that limit ordinals are precisely the ordinals of the form $L(I)$.

For an ordinal system $\mathcal{O}$ relative to $\mathfrak{U}$, consider the limit-successor system $(\mathcal{O}, \bigvee, \_^+)$, where $\bigvee$ is the usual join restricted to the domain required by (C1) (i.e. successor-closed elements of $\mathsf{P}_\mathfrak{U} X$).

Step 5: *We show that (C1–5) hold for the limit-successor system $(\mathcal{O}, \bigvee, \_^+)$ and that the corresponding specialisation preorder matches with the order of $\mathcal{O}$. In this step we also show that $\overline{I} = \{\forall\!\!\!/\, I\}^>$ holds for each $I \in \mathsf{P}_\mathfrak{U}\mathcal{O}$.*

By Theorem 53, L is indeed defined over the entire domain required in (C1). To prove (C2), first we establish that

$$\overline{I} = \{\forall\!\!\!/\, I\}^>$$

for each $I \in \mathsf{P}_\mathfrak{U}\mathcal{O}$. It is easy to see that $\{\forall\!\!\!/\, I\}^>$ is closed, so $\overline{I} \subseteq \{\forall\!\!\!/\, I\}^>$. To show $\{\forall\!\!\!/\, I\}^> \subseteq \overline{I}$, let $x \in \{\forall\!\!\!/\, I\}^>$. We have well-ordering and hence total order by Theorem 50. Then $x < \forall\!\!\!/\, I$ and so $x \leqslant y \in I$ for some y. Consider

$$y' = \min\{y \in \overline{I} \mid x \leqslant y\}.$$

We consider two cases:

Case 1: y' is a successor ordinal. Then $y' = y''^+$ for some ordinal y''. Since $y' \in \overline{I}$, we must have $y'' \in \overline{I}$. Then $y'' < x$ and so $y' \leqslant x$ by (L3). This gives $x = y'$ and so $x \in \overline{I}$.

Case 2: y' is a limit ordinal. Then $\{y'\}^>$ is successor-closed. Furthermore, we have

$$y' = \forall\!\!\!\bigvee\{y'\}^> = \bigvee(\{y'\}^>)^+ = \bigvee\{y'\}^>.$$

By closure of $\overline{I}$, we get $\{y'\}^> \subseteq \overline{I}$. Since $y < x$ for all $y < y'$, we get $y' \leqslant x$. This gives $x = y'$ and so $x \in \overline{I}$.

We have thus established that the equality $\overline{I} = \{\forall\!\!\!\bigvee I\}^>$ holds for each $I \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$. From this it follows that the specialisation preorder matches with the order of $\mathcal{O}$. We then get that (C2) holds by the fact that if down-closures of two subsets of a poset are equal, then so are their joins. Thus $(\mathcal{O}, \bigvee, \_^+)$ is a $\mathfrak{U}$-counting system.

It remains to show that (C3–5) hold. Consider a successor-closed $I \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$. I has no maximum element thanks to (L1) and thus, $\bigvee I = \forall\!\!\!\bigvee I$ by Lemma 46. By the same lemma, $\forall\!\!\!\bigvee I$ cannot be a successor if I has no maximum. Thus, $(\_^+)^{-1}\{\bigvee I\} = \varnothing$, which is the first part of (C3). Since $(\bigvee I)^+ > \bigvee I \geqslant x$ for each $x \in I$, we have that $(\bigvee I)^+ \notin I$, which means the second part of (C3) also holds, i.e. $I \cap \{(\bigvee I)^+\} = \varnothing$.

We already know that $\_^+$ is injective, so to see that (C4) holds, consider another successor-closed $J \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$. If $\bigvee I = \bigvee J$, then

$$\overline{I} = \{\forall\!\!\!\bigvee I\}^> = \{\bigvee I\}^> = \{\bigvee J\}^> = \{\forall\!\!\!\bigvee J\}^> = \overline{J}.$$

Thus (C4) holds. Finally, consider a successor-closed subset J of $\mathcal{O}$ where $\bigvee I \in J$ for all successor-closed subsets I of J such that $I \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$. Then $J = \mathcal{O}$ if it satisfies (I1) and (I2) in our formulation of transfinite induction. We check both:

I1. $J^+ \subseteq J$ follows from the fact that J is successor-closed.

I2. Let x be a limit ordinal such that $\{x\}^> \subseteq J$. Then $x = \bigvee\{x\}^> \in J$ by Lemma 52.

Since both of these conditions hold, we can conclude that $J = \mathcal{O}$, and thus (C5) holds. This completes the proof. ∎

Definition 79 (ordinal $\mathfrak{U}$-counting system)**.** We call a $\mathfrak{U}$-counting system $(\mathcal{O}, \bigvee, \_^+)$ an *ordinal $\mathfrak{U}$-counting* system when conditions (C3–5) hold, i.e. when $(\mathcal{O}, \leqslant)$ is an ordinal system relative to $\mathfrak{U}$.

3.3 The universal property

In this section we establish that ordinal $\mathfrak{U}$-counting systems are the initial objects in appropriate categories of $\mathfrak{U}$-counting systems.

Definition 80 (morphism of $\mathfrak{U}$-counting systems)**.** Given $\mathfrak{U}$-counting systems (X_1, L_1, s_1) and (X_2, L_2, s_2), a *morphism of $\mathfrak{U}$-counting systems* is a function $f\colon X_1 \to X_2$ such that

(a) $f \circ s_1 = s_2 \circ f$;

(b) $f(L_1(I)) = L_2(fI)$ for any successor-closed $I \in \mathsf{P}_{\mathfrak{U}} X_1$.

Note that by (a), both sides of the equality in (b) are defined for all successor closed $I \in \mathsf{P}_{\mathfrak{U}} X_1$.

Lemma 81. *A morphism $f\colon (X_1, L_1, s_1) \to (X_2, L_2, s_2)$ is an increasing function under the specialisation preorders of X_1 and X_2.*

Proof. Consider $x, y \in X_1$ such that $x \leqslant y$. We consider two cases.

Case 1: $x \leqslant_{s_1} y$. Then ${s_1}^k(x) = y$ for some $k \in \mathbb{N}$. Then, by Definition 80,

$$\begin{aligned} f(y) &= f({s_1}^k(x)) \\ &= {s_2}^k(f(x)), \end{aligned}$$

and thus $f(x) \leqslant_{s_2} f(y)$.

Case 2: $x \leqslant_{L_1} z \leqslant_{s_1} y$ for some $z \in X_1$. Then there exists a successor-closed $I \in \mathsf{P}_{\mathfrak{U}} X_1$ such that $x \in I$ and $L_1(I) = z$. But then

$$\begin{aligned} f(z) &= f(L_1(I)) \\ &= L_2(fI), \end{aligned}$$

and since $f(x) \in fI$, this means $f(x) \leqslant_{L_2} f(z)$. Since $f(z) \leqslant_{s_2} f(y)$ by the argument in Case 1, this gives us $f(x) \leqslant_{L_2} f(z) \leqslant_{s_2} f(y)$

Either way, $f(x) \leqslant f(y)$, and thus f is increasing. ∎

Lemma 82. *For a fixed universe $\mathfrak{U}$, $\mathfrak{U}$-counting systems and the morphisms between them form a category under the usual composition of functions. Isomorphisms in this category are bijections between $\mathfrak{U}$-counting systems which preserve both succession and limiting.*

Proof. Consider $\mathfrak{U}$-counting systems (X_1, L_1, s_1), (X_2, L_2, s_2), (X_3, L_3, s_3) and (X_4, L_4, s_4), and functions $f\colon X_1 \to X_2$, $g\colon X_2 \to X_3$ and $h\colon X_3 \to X_4$ that satisfy (a–b) in Definition 80. Identity maps clearly satisfy (a–b), and $h \circ (g \circ f) = (h \circ g) \circ f$ follows from the fact that f, g and h are set functions. It remains to prove that $g \circ f$ also satisfies (a–b).

a. $(g \circ f) \circ s_1 = s_3 \circ (g \circ f)$ follows from associativity of set functions and (a) in Definition 80.

b. $(g \circ f)(L_1(I)) = g(L_2(fI)) = L_3((g \circ f)I)$.

This proves that $\mathfrak{U}$-counting systems do indeed form a category with morphisms as described in Definition 80.

Now suppose f is a bijection, i.e. a function $f^{-1}\colon X_2 \to X_1$ exists such that $f \circ f^{-1} = 1_Y$ and $f^{-1} \circ f = 1_X$. We confirm that f^{-1} satisfies (a–b) in Definition 80:

a.
$$\begin{aligned} & f \circ s_1 = s_2 \circ f \\ \Rightarrow \quad & f^{-1} \circ f \circ s_1 \circ f^{-1} = f^{-1} \circ s_2 \circ f \circ f^{-1} \\ \Rightarrow \quad & s_1 \circ f^{-1} = f^{-1} \circ s_2. \end{aligned}$$

b. Consider a successor-closed $I \in \mathsf{P}_{\mathfrak{U}} X_2$. As we just established, f^{-1} preserves successors and thus $f^{-1}I$ is also successor-closed. Then $L_1(f^{-1}I)$ is defined, and thus, using (b) in Definition 80,

$$\begin{aligned} & f(L_1(f^{-1}I)) = L_2(ff^{-1}I) \\ \Rightarrow \quad & f^{-1}f(L_1(f^{-1}I)) = f^{-1}(L_2(I)) \\ \Rightarrow \quad & L_1(f^{-1}I) = f^{-1}(L_2(I)). \end{aligned}$$

Thus f^{-1} is a morphism of $\mathfrak{U}$-counting systems, and we can conclude that f is indeed an isomorphism of $\mathfrak{U}$-counting systems. ∎

Clearly, the property of being an ordinal $\mathfrak{U}$-counting system is stable under isomorphism of $\mathfrak{U}$-counting systems. By Theorem 59 and Theorem 78, an ordinal $\mathfrak{U}$-counting system exists and is given by the $\mathfrak{U}$-ordinals. We will now see that ordinal $\mathfrak{U}$-counting systems are precisely the initial objects in the category of $\mathfrak{U}$-counting systems.

Theorem 83. *For any universe $\mathfrak{U}$, a $\mathfrak{U}$-counting system is an initial object in the category of $\mathfrak{U}$-counting systems if and only if it is an ordinal $\mathfrak{U}$-counting system.*

Proof. Since we know that an ordinal $\mathfrak{U}$-counting system exists (Theorem 59), and that the property of being an ordinal $\mathfrak{U}$-counting system is stable under isomorphism, it suffices to show that any ordinal $\mathfrak{U}$-counting system is an initial object in the category of $\mathfrak{U}$-counting systems. By Theorem 78, an ordinal $\mathfrak{U}$-counting system has the form $(\mathcal{O}, \bigvee, \_^+)$, where $\mathcal{O}$ is an ordinal system relative to $\mathfrak{U}$ and $\bigvee$ is the join defined for exactly the successor-closed subsets $I \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$ in the $\mathfrak{U}$-counting system.

For any $\mathfrak{U}$-counting system (X, L, s), if a morphism $(\mathcal{O}, \bigvee, \_^+) \to (X, L, s)$ exists, it must be the unique function f defined by the transfinite recursion

(i) $f(x^+) = s(f(x))$ for any $x \in \mathcal{O}$;

(ii) $f(x) = f(\bigvee\{y \mid y < x\}) = L(\{f(y) \mid y < x\})$ for any limit ordinal x (Lemma 52).

We now prove that the function f defined by the recursion above is a morphism. It preserves succession by (i). Consider $I \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$ closed under successors. Then $\bigvee I$ is a limit ordinal and $\overline{I} = \{\bigvee I\}^>$, by Theorem 78. By definition of f, we then have

$$f(\bigvee I) = L(\{f(y) \mid y < \bigvee I\}) = L(f\overline{I}).$$

We will now prove $\overline{f\overline{I}} = \overline{fI}$. We clearly have $fI \subseteq \overline{f\overline{I}}$, so it suffices to show that $f\overline{I} \subseteq \overline{fI}$. This is equivalent to showing $\overline{I} \subseteq f^{-1}\overline{fI}$, which would follow if we prove $f^{-1}\overline{fI}$ is closed. If $x^+ \in f^{-1}\overline{fI}$, then $s(f(x)) = f(x^+) \in \overline{fI}$. Therefore, $f(x) \in \overline{fI}$ and so $x \in f^{-1}\overline{fI}$. If $\bigvee J \in f^{-1}\overline{fI}$, then (as $\bigvee J$ is a limit ordinal by Theorem 78)

$$L(\{f(y) \mid y < \bigvee J\}) = f(\bigvee J) \in \overline{fI},$$

which implies $\{f(y) \mid y < \bigvee J\} \subseteq \overline{fI}$. This gives $J \subseteq \{\bigvee J\}^> \subseteq f^{-1}\overline{fI}$. Note that we have the first of these two subset inclusions due to the fact that $\bigvee J = \forall J$ thanks to Theorem 78. This proves that $f^{-1}\overline{fI}$ is closed. So $\overline{f\overline{I}} = \overline{fI}$. We therefore get $f(\bigvee I) = L(f\overline{I}) = L(fI)$, showing that f is indeed a morphism $(\mathcal{O}, \bigvee, \_^+) \to (X, L, s)$. ■

Example 84 (natural number system)**.** Consider the case where every element in $\mathfrak{U} \neq \varnothing$ is a finite set. Then each triple $(X, 0, s)$, where X is a set, s is a function $s\colon X \to X$, and $0 \in X$, can be seen as a $\mathfrak{U}$-counting system for the same s, with $L(I) = 0$ for each finite successor-closed I. A morphism $f\colon (X_1, 0_1, s_1) \to (X_2, 0_2, s_2)$ between such $\mathfrak{U}$-counting systems is a function $f\colon X_1 \to X_2$ such that $s_2 f = f s_1$ and $f(0_1) = 0_2$. The natural number system $(\mathbb{N}, 0, s)$, with its usual successor function $s(n) = n + 1$, is an initial object in the category of such $\mathfrak{U}$-counting systems by Theorem 22. Theorem 83 presents the natural number system as an initial object in the category of all $\mathfrak{U}$-counting systems. It is not surprising that the natural number system is initial in this larger category too, since the empty set is the only finite successor-closed subset of $\mathbb{N}$.

Chapter 4

Final remarks

In this final chapter of the thesis we indicate some directions of further research on the theme of this thesis.

A feature of the abstract theory of ordinal number systems is that the ordinals themselves are less reliant on the structure of the element relation. Additionally, while we often use Zermelo-Fraenkel axioms for $\mathfrak{U}$-small and $\mathfrak{U}$-moderate sets of ordinals, we are much more frugal with these axioms when it comes to $\mathfrak{U}$-large sets. With these two remarks in mind, our next move is clear: further generalisation. The ultimate goal would be to export our theory to a category-theoretic context, such as that of topos theory, where we have neither the full power of the element relation nor the full extent of the Zermelo-Fraenkel axioms. In Section 4.1 we outline only a first step in this direction, which is to identify which Zermelo-Fraenkel axioms we have used so far and whether we have used them within the bounds of our universe or in their general form. We call this variation on Zermelo-Fraenkel set theory a 'relative' set theory.

Section 4.2 gives a very short outline of an alternative approach to set theory developed by André Joyal and Ieke Moerdijk in [7], which is partially based on topos theory and where known universal properties of the ordinal number system first appear. This could be a suitable framework for the generalisation we have in mind, especially given that our first universal property of the ordinal number system already features in algebraic set theory as a way of introducing abstract ordinal systems.

In Section 4.3 we give the first steps towards further developing the theory of ordinal numbers in our context. We define a 'cumulative hierarchy' of well-founded sets relative to our universe, first in the traditional style, and then using the universal properties from Chapter 2 and Chapter 3. We use the cumulative hierarchy as a context to define a strictly associative 'addition' of sets, which was the original inspiration that led to the research contained in this thesis.

4.1 Relative set theory

The first-order language of 'relative set theory' proposed in this section extends the first-order language of Zermelo-Fraenkel set theory with a constant term $\mathfrak{U}$, which we call a *universe*. We will refer to elements of $\mathfrak{U}$ as *small* sets and replace some of the Zermelo-Fraenkel axioms with variations relative to $\mathfrak{U}$, which apply specifically to small sets.

For small sets, we include the following 'axiom of transitivity', which matches (U1). While there is no Zermelo-Fraenkel axiom of transitivity, it is built into the first-order language of set theory – since all terms in that language are sets, the statement $x \in y$ is a claim about the sets x and y, which are each, per definition, members of the Zermelo-Fraenkel universe.

Axiom 85 (transitivity)**.** If $x \in y$ and y is a small set, then x is a small set.

The axiom of extensionality and the axiom schema of restricted comprehension are both needed in their original form. One can think of them as analogues of class equivalence and class comprehension in Zermelo-Fraenkel set theory.

Axiom 86 (extensionality)**.** For all sets X and Y, if $x \in X \Leftrightarrow x \in Y$ for all x, then $X = Y$.

Axiom schema 87 (restricted comprehension)**.** If $\varphi(x, v_1, \ldots, v_n)$ is a formula with free variables $x, v_1, \ldots, v_n$, then for any set X and arbitrary sets $v_1, \ldots, v_n$, there exists a set

$$Y = \{x \in X \mid \varphi(x, v_1, \ldots, v_n)\}.$$

Our universe of small sets needs to be closed under pairing, as stated in (U2), but the axiom of relative pairing is never used outside the universe. Our pairing axiom is thus relative to the universe of small sets.

Axiom 88 (relative pairing)**.** If x and y are small sets, then there exists a small set that has x and y as elements.

The axiom of power set is not required for small sets, since we did not assume (U4) for general universes. Since $\mathsf{P}_{\mathfrak{U}}\mathfrak{U} = \mathfrak{U}$ (Lemma 36), we do not require an axiom of power set at all.

The union operation is required within $\mathfrak{U}$, but is never used outside it. Together with the axiom schema of relative replacement, it corresponds to (U3).

Axiom 89 (relative union)**.** For every small set X, there exists a small set

$$\bigcup X = \{y \mid \exists_x [y \in x \wedge x \in X]\},$$

called the *union* of X.

If we want our universe to be nonempty we need to include an axiom of the empty set relative to the universe.

Axiom 90 (relative empty set)**.** There exists a small set $\varnothing$ with no elements.

Axiom 91 (relative choice)**.** If X is a small set of mutually disjoint sets, then there exists a small subset of $\bigcup X$ that has exactly one element in common with each element of X.

Axiom schema 92 (relative replacement)**.** If $f\colon X \to \mathfrak{U}$ is a function and $Y \subseteq X$ is a small set, then fY is a small set.

4.2 Algebraic set theory

We usually think of sets as collections of elements, as Cantor conceived them, and as Zermelo, Fraenkel, and others formalised with axioms on the membership relation $\in$. Indeed, this concrete approach is followed throughout this thesis and in every cited source but one – the exposition of 'algebraic set theory' by André Joyal and Ieke Moerdijk in 1995 [7].

Joyal and Moerdijk take a category-theoretic approach, where the starting point is an abstraction of the category of sets, fulfilling axioms that to a certain extent imitate Zermelo-Fraenkel behaviour of this more general axiomatic framework. They further endow this context with a distinguished class of 'small' maps, which provides a way of fixing a universe in their framework. Specifically, in the context considered in this thesis, we could define small maps as functions whose fibres are small sets in the sense of Section 1.5. Verification that our context fulfils the axioms of small sets presented in [7] is beyond the scope of this thesis; however, it should follow easily from the discussion of examples in Chapter IV of [7].

A key notion in [7] is that of a 'Zermelo-Fraenkel algebra' (or 'ZF algebra'). Each ZF algebra is partially ordered and has a successor operation and a join operation. Morphisms between ZF algebras are small maps which preserve both successors and joins. The 'collecting systems', which we introduced in Section 2.5, are, in fact, ZF algebras with monotone successors, once we specialise their context to ours. In that section we identify our abstract ordinal systems as initial objects in categories of collecting systems whose morphisms preserve incremented joins (this is slightly more general than the morphisms in [7], which preserve both joins and successors – see Remark 65). For Joyal and Moerdijk, this universal property is one of the definitions of an ordinal system.

4.3 Associative set addition

Throughout this section we work with a fixed universe $\mathfrak{U}$ satisfying (U4), and a fixed ordinal system $(\mathcal{O}, \leqslant)$ relative to $\mathfrak{U}$ (for example, $(\mathcal{O}, \leqslant)$ could be the concrete ordinal system from Section 2.4).

By applying transfinite recursion (Theorem 57) for the triple $(\mathfrak{U}, u, p)$, where u and p are defined by $u(x) = \bigcup x$ and $p(X) = \mathsf{P}X$, we get a function

$$\mathcal{O} \to \mathfrak{U}, \quad \alpha \mapsto \mathcal{V}_\alpha,$$

satisfying the following:

$$\mathcal{V}_{\alpha^+} = \mathsf{P}\mathcal{V}_\alpha; \qquad \text{(for each ordinal } \alpha)$$

$$\mathcal{V}_\gamma = \bigcup\{\mathcal{V}_\beta \mid \beta < \gamma\}. \qquad \text{(for each limit ordinal } \gamma)$$

We further define

$$\mathfrak{V} = \bigcup\{\mathcal{V}_\beta \mid \beta \in \mathcal{O}\},$$

where $\{\mathcal{V}_\beta \mid \beta \in \mathcal{O}\}$ is called the set of *stages* of $\mathfrak{V}$. For each element $x \in \mathfrak{V}$, the *rank* of x is defined as

$$\mathsf{rank}\, x = \min\{\beta \mid x \subseteq \mathcal{V}_\beta\}.$$

This defines a function $\mathsf{rank}\colon \mathfrak{V} \to \mathcal{O}$.

$\mathfrak{V}$ is nothing other than the usual *cumulative hierarchy of well-founded sets* (see e.g. [29]), relativised to our universe $\mathfrak{U}$. We recall the definition of a well-founded set and prove that all elements of $\mathfrak{V}$ are indeed well-founded further below.

Lemma 93. *For each ordinal α, we have $\mathcal{V}_\alpha \subset \mathcal{V}_{\alpha^+}$.*

Proof. We prove this by transfinite induction.

Successor case: Suppose that for some ordinal α it holds that $\mathcal{V}_\alpha \subset \mathcal{V}_{\alpha^+}$. Then

$$\mathcal{V}_{\alpha^+} = \mathsf{P}\mathcal{V}_\alpha \subset \mathsf{P}\mathcal{V}_{\alpha^+} = \mathcal{V}_{\alpha^{++}}.$$

Limit case: Suppose that for some limit ordinal γ and each $\beta < \gamma$ it holds that $\mathcal{V}_\beta \subset \mathcal{V}_{\beta^+}$. Consider $x \in \mathcal{V}_\gamma$. Then $x \in \mathcal{V}_\beta$ for some $\beta < \gamma$. Then, by induction assumption, $x \in \mathcal{V}_{\beta^+} = \mathsf{P}\mathcal{V}_\beta$, and thus $x \subseteq \mathcal{V}_\beta \subset \mathcal{V}_\gamma$. This gives us $x \in \mathsf{P}\mathcal{V}_\gamma = \mathcal{V}_{\gamma^+}$, and thus $\mathcal{V}_\gamma \subseteq \mathcal{V}_{\gamma^+}$. Since $\mathcal{V}_\gamma \neq \mathcal{V}_{\gamma^+}$, we now have $\mathcal{V}_\gamma \subset \mathcal{V}_{\gamma^+}$.

We can conclude that $\mathcal{V}_\alpha \subset \mathcal{V}_{\alpha^+}$ holds for all ordinals α. ∎

Lemma 94. *$\mathcal{V}_\beta \subset \mathcal{V}_\alpha$ if and only if $\beta < \alpha$.*

Proof. We prove this by transfinite induction.

Successor case: Suppose for some α and all $\beta < \alpha$ it holds that $\mathcal{V}_\beta \subset \mathcal{V}_\alpha$. Then by Lemma 93, $\mathcal{V}_\beta \subset \mathcal{V}_\alpha \subset \mathcal{V}_{\alpha^+}$.

Limit case: Consider a limit ordinal γ, and suppose that for all $\beta < \gamma$ it holds that $\mathcal{V}_\delta \subset \mathcal{V}_\beta$ whenever $\delta < \beta$. Since $\mathcal{V}_\gamma = \bigcup\{\mathcal{V}_\beta \mid \beta < \gamma\}$, we have $\mathcal{V}_\beta \subseteq \mathcal{V}_\gamma$ for all $\beta < \gamma$. Since γ is a limit ordinal, we also have $\mathcal{V}_{\beta^+} \subseteq \mathcal{V}_\gamma$ for all $\beta < \gamma$, and by Lemma 93, this gives us $\mathcal{V}_\beta \subset \mathcal{V}_{\beta^+} \subseteq \mathcal{V}_\gamma$.

This proves that $\mathcal{V}_\beta \subset \mathcal{V}_\alpha$ whenever $\beta < \alpha$. To see the inverse, suppose that $V_\beta \subset V_\alpha$. Then $V_\alpha \neq V_\beta$, and also $V_\alpha \not\subset V_\beta$. Thus, $\alpha \neq \beta$, and by contrapositive, $\alpha \not< \beta$. Thus $\alpha \nleqslant \beta$, or equivalently, $\beta < \alpha$. We can conclude that $\mathcal{V}_\beta \subset \mathcal{V}_\alpha$ if and only if $\beta < \alpha$. ∎

Lemma 95. *For any $S \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$, it holds that*

$$\bigcup\{\mathcal{V}_\alpha \mid \alpha \in S\} = \bigcup\{\mathcal{V}_\beta \mid \beta < \dot{\bigvee} S\} = \mathcal{V}_{\bigvee S}.$$

Proof. Since $S \subseteq \{\beta \mid \beta < \dot{\bigvee} S\}$, we know that

$$\bigcup\{\mathcal{V}_\alpha \mid \alpha \in S\} \subseteq \bigcup\{\mathcal{V}_\beta \mid \beta < \dot{\bigvee} S\}.$$

For the other direction, suppose $x \in \bigcup\{\mathcal{V}_\beta \mid \beta < \dot{\bigvee} S\}$. Then for some $\beta < \dot{\bigvee} S$, we have $x \in \mathcal{V}_\beta$. Since $\beta < \dot{\bigvee} S$, there must exist $\alpha \in S$ such that $\beta \leqslant \alpha$, or else we would have $\beta \in S^<$, and thus $\dot{\bigvee} S = \min S^< \leqslant \beta$. Then $x \in \mathcal{V}_\beta \subseteq \mathcal{V}_\alpha$ by Lemma 94, and thus $x \in \bigcup\{\mathcal{V}_\alpha \mid \alpha \in S\}$, giving us

$$\bigcup\{\mathcal{V}_\beta \mid \beta < \dot{\bigvee} S\} \subseteq \bigcup\{\mathcal{V}_\alpha \mid \alpha \in S\}.$$

We can conclude that

$$\bigcup\{\mathcal{V}_\alpha \mid \alpha \in S\} = \bigcup\{\mathcal{V}_\beta \mid \beta < \dot{\bigvee} S\}.$$

Now, since $\mathcal{V}_\beta \subset \mathcal{V}_\alpha$ whenever $\beta < \alpha$, if S has a maximum element α, then

$$\bigcup\{\mathcal{V}_\beta \mid \beta \in S\} = \mathcal{V}_\alpha = \mathcal{V}_{\bigvee S}.$$

If S does not have a maximum element, then $\bigvee S = \gamma$ for some limit ordinal γ. Then

$$\bigcup\{\mathcal{V}_\beta \mid \beta < \dot{\bigvee} S\} = \bigcup\{\mathcal{V}_\beta \mid \beta < \gamma\} = \mathcal{V}_\gamma.$$

Either way, we can conclude, as desired, that

$$\bigcup\{\mathcal{V}_\alpha \mid \alpha \in S\} = \bigcup\{\mathcal{V}_\beta \mid \beta < \dot{\bigvee} S\} = \mathcal{V}_{\bigvee S}.$$ ∎

Theorem 96. *The set X of stages of $\mathfrak{V}$, ordered by subset inclusion, is order-isomorphic to $(\mathcal{O}, \leqslant)$. The join of a set $S \in \mathsf{P}_{\mathfrak{U}}X$ of stages is given by $\bigvee S = \bigcup S$, and the successor of a stage $\mathcal{V}_\alpha$ by $\mathsf{P}\mathcal{V}_\alpha = \mathcal{V}_{\alpha^+}$. Thus, $(X, \subseteq)$ can be seen as an ordinal system, in which the incremented join of a set $S \in \mathsf{P}_{\mathfrak{U}}X$ of stages is given by*

$$\overset{+}{\bigvee} S = \bigcup \{\mathsf{P}\mathcal{V}_\beta \mid \mathcal{V}_\beta \in S\}.$$

Proof. Let $X = \{\mathcal{V}_\alpha \mid \alpha \in \mathcal{O}\}$. The function $f \colon \mathcal{O} \to X$ defined by

$$\alpha \mapsto \mathcal{V}_\alpha$$

is clearly surjective, and thus, by Lemma 94, it is an order isomorphism (see Definition 21) between the posets $(\mathcal{O}, \leqslant)$ and $(X, \subseteq)$. From this order isomorphism we get, for each ordinal α, that $\mathcal{V}_{\alpha^+} = \mathsf{P}\mathcal{V}_\alpha$ is the successor of $\mathcal{V}_\alpha$, and thus the successor function of $(X, \subseteq)$ is given by the function $p \colon X \to X$ defined by $p(\mathcal{V}_\alpha) = \mathsf{P}\mathcal{V}_\alpha$.

Now let $S \in \mathsf{P}_{\mathfrak{U}}X$. From Lemma 94, we know that $\bigcup S \in X$. Then by Lemma 45,

$$\overset{+}{\bigvee} S = \bigvee pS = \bigcup \{\mathsf{P}x \mid x \in S\}. \qquad \blacksquare$$

Theorem 97. *$(\mathfrak{U}, p, \subseteq)$ is a $\mathfrak{U}$-collecting system, where $p(x) = \mathsf{P}x$ for each $x \in \mathfrak{U}$, and the collecting function $\overset{+}{\bigvee} \colon \mathsf{P}_{\mathfrak{U}}X \to X$ is given by*

$$S \mapsto \bigcup \{\mathsf{P}x \mid x \in S\}.$$

Proof. Notice that, since $\mathfrak{U}$ satisfies (U4), $\mathsf{P}x \in \mathfrak{U}$ for each $x \in \mathfrak{U}$, and thus p defines a function $\mathfrak{U} \to \mathfrak{U}$. We prove that $(\mathfrak{U}, \subseteq)$ satisfies (CA1–2).

CA1. Let $S \in \mathsf{P}_{\mathfrak{U}}\mathfrak{U}$. By Lemma 36, $\mathsf{P}_{\mathfrak{U}}\mathfrak{U} = \mathfrak{U}$ and thus $S \in \mathfrak{U}$. Therefore, by Theorem 27 (b), $\bigcup S \in \mathfrak{U}$, and since $\bigcup S$ is per definition the smallest set that has each element of S as a subset, this gives us $\bigvee S = \bigcup S$.

CA2. Consider $x, y \in \mathfrak{U}$ such that $x \subseteq y$. Then $p(x) = \mathsf{P}x \subseteq \mathsf{P}y = p(y)$, and thus p is increasing in $(\mathfrak{U}, \subseteq)$.

So $(\mathfrak{U}, p, \subseteq)$ is a $\mathfrak{U}$-collecting system where $p(x) = \mathsf{P}x$ for each $x \in \mathfrak{U}$, and $\bigcup S$ is the join of S for each $S \in \mathsf{P}_{\mathfrak{U}}\mathfrak{U}$. It follows that, for each $S \in \mathsf{P}_{\mathfrak{U}}\mathfrak{U}$,

$$\overset{p}{\bigvee} S = \bigcup \{\mathsf{P}x \mid x \in S\}. \qquad \blacksquare$$

The *transitive closure* of a set X is a transitive set X^T such that $X \subseteq X^T$, and if T is any transitive set such that $X \subseteq T$, then $X^T \subseteq T$, i.e. X^T is the intersection of all such sets T (provided such a set T exists, which we will prove in Lemma 98).

Lemma 98. *The transitive closure of a set X can be constructed by setting*

$$X^T = \bigcup\{\bigcup^n X \mid n \in \mathbb{N}\},$$

where $\bigcup^n$ is defined recursively by

$$\bigcup^0 X = X;$$
$$\bigcup^{n+1} X = \bigcup\bigcup^n X.$$

Furthermore, if each element of X is transitive, then $X^T = X \cup \bigcup X$.

Proof. Let $X^T = \bigcup\{\bigcup^n X \mid n \in \mathbb{N}\}$. To prove X^T is transitive, consider any $x \in X^T$. Then $x \in \bigcup^n X$ for some $n \in \mathbb{N}$, and thus

$$x \subseteq \bigcup\bigcup^n X = \bigcup^{n+1} X \subseteq X^T.$$

So X^T is transitive. Now consider any transitive set T such that $X \subseteq T$. Suppose that for some $n \in \mathbb{N}$ it holds that $\bigcup^n X \subseteq T$. Then, since T is transitive, $\bigcup\bigcup^n X \subseteq T$. By induction, this gives us $\bigcup^n \subseteq T$ for all $n \in \mathbb{N}$. We can conclude from this that $X^T = \bigcup\{\bigcup^n X \mid n \in \mathbb{N}\} \subseteq T$, and thus X^T is the transitive closure of X.

Now suppose each element of X is transitive. Naturally, $X \cup \bigcup X \subseteq X^T$, so let us prove that $X \cup \bigcup X$ is transitive. Consider any $x \in X \cup \bigcup X$. If $x \in X$, then $x \subseteq \bigcup X \subseteq X \cup \bigcup X$. Otherwise, $x \in \bigcup X$, and thus $x \in y$ for some $y \in X$. Since y is transitive, this gives us $x \subseteq y \subseteq \bigcup X \subseteq X \cup \bigcup X$. Either way, $x \subseteq X \cup \bigcup X$, and thus $X \cup \bigcup X$ is transitive. Then it is a transitive subset of X^T, and hence equal to X^T. ■

Lemma 99. *If $\mathfrak{U}$ is empty or has an infinite element and $X \in \mathfrak{U}$, then $X^T \in \mathfrak{U}$.*

Proof. If $\mathfrak{U}$ is empty, this is trivial. Suppose $\mathfrak{U}$ has an infinite element. Then $\omega \in \mathfrak{U}$ (Lemma 29). Since $\bigcup^n X \in \mathfrak{U}$ for each $n \in \omega$ (induction using Theorem 27 (b)), we can define a function $f\colon \omega \to \mathfrak{U}$ by $n \mapsto \bigcup^n X$. Then $X^T = \bigcup\{\bigcup^n X \mid n \in \omega\} = \bigcup f\omega \in \mathfrak{U}$ by (U3) and Lemma 98. ■

Theorem 100. *Let $X = \{x \in \mathfrak{U} \mid x \subseteq \mathsf{P}x\}$ be the set of transitive elements of $\mathfrak{U}$. Then (X, u, p) is a $\mathfrak{U}$-counting system, where $p(x) = \mathsf{P}x$ for each $x \in X$, and $u(I) = \bigcup I$ for each successor-closed $I \in \mathsf{P}_{\mathfrak{U}}X$.*

Proof. Notice that for any $x \in X$, we have $\mathsf{P}x \subseteq \mathsf{PP}x$, and thus we can define a function $p\colon X \to X$ such that $p(x) = \mathsf{P}x$ for each $x \in X$. Now consider a set $I \in \mathsf{P}_{\mathfrak{U}}X$ that is closed under p. Since I is a set of transitive sets, $\bigcup I$ is transitive, and $\bigcup I \in \mathfrak{U}$ by Theorem 27. Thus $\bigcup I \in X$, and we can define a function $u\colon \{I \in \mathsf{P}_{\mathfrak{U}}X \mid pI \subseteq I\} \to X$

by $I \mapsto \bigcup I$. Clearly, (X, u, p) is a limit-successor system that satisfies (C1). We must prove that it satisfies (C2).

Let $I \in \mathsf{P}_{\mathfrak{U}}X$ be successor-closed. We will prove that $\bigcup I = \bigcup \bar{I}$ by showing that $\bigcup I \subseteq \bigcup \bar{I} \subseteq \bigcup I^T$ and $\bigcup I = \bigcup I^T$ both hold, where I^T is the transitive closure of I.

First, let us prove that any transitive, successor-closed $J \in \mathsf{P}_{\mathfrak{U}}X$ is closed in the limit-successor system topology on X (see Definition 70). We need to prove that that $p^{-1}J \subseteq J$ and $\bigcup u^{-1}J \subseteq J$. Let

$$\begin{aligned} x \in p^{-1}J &= \{x \in X \mid p(x) \in J\} \\ &= \{x \in X \mid \mathsf{P}x \in J\}. \end{aligned}$$

So $\mathsf{P}x \in J$, and since J is transitive and $x \in \mathsf{P}x$, this gives us $x \in J$. Now let

$$\begin{aligned} x \in \bigcup u^{-1}J &= \bigcup\{A \in \operatorname{dom} u \mid u(A) \in J\} \\ &= \bigcup\{A \in \mathsf{P}_{\mathfrak{U}}X \mid [pA \subseteq A] \wedge [\bigcup A \in J]\}. \end{aligned}$$

Then $x \in A$ for some A that satisfies both $pA \subseteq A$ and $\bigcup A \in J$, and thus $\mathsf{P}x = p(x) \in A$. Also, since J is transitive, $\bigcup A \subseteq J$, giving us the following:

$$x \in \mathsf{P}x \subseteq \bigcup A \subseteq J,$$

giving us $x \in J$. We can conclude that $\bigcup u^{-1}J \subseteq J$. Since $p^{-1}J \subseteq J$ and $\bigcup u^{-1}J \subseteq J$ both hold, J is closed in (X, u, p).

Now consider any successor-closed $I \in \mathsf{P}_{\mathfrak{U}}X$. The transitive closure I^T of I can be computed as $I^T = I \cup \bigcup I$ by Lemma 98. To prove that I^T is successor-closed, consider any $x \in I^T$. Then either $x \in I$, or $x \in \bigcup I$. If $x \in I$, then $p(x) \in I \subseteq I^T$, since I is successor-closed. If $x \in \bigcup I$, then $x \in y$ for some $y \in I$, and thus $x \subseteq y$. This gives us $\mathsf{P}x \subseteq \mathsf{P}y$, i.e. $\mathsf{P}x \in \mathsf{PP}y$. Since I is successor-closed, $\mathsf{PP}y = p(p(y)) \in I \subseteq I^T$, and thus, by transitivity of I^T,

$$p(x) = \mathsf{P}x \in \mathsf{PP}y \subseteq I^T.$$

We can conclude that I^T is a successor-closed, transitive element of $\mathsf{P}_{\mathfrak{U}}X$, and, as such, is closed in the limit-successor system topology on X. This leads us to conclude that

$$I \subseteq \bar{I} \subseteq I^T.$$

Now, note that as a union of transitive sets, $\bigcup I$ is transitive and thus $\bigcup\bigcup I \subseteq \bigcup I$. Then

$$\begin{aligned} \bigcup I^T &= \bigcup(I \cup \bigcup I) \\ &= \bigcup I \cup \bigcup\bigcup I \\ &= \bigcup I. \end{aligned}$$

From these two identities, we can squeeze out the following:

$$\bigcup I = \bigcup \bar{I} = \bigcup I^T.$$

Thus, for all I, J in the domain of u, if $\bar{I} = \bar{J}$, then

$$u(I) = \bigcup I = \bigcup \bar{I} = \bigcup \bar{J} = \bigcup J = u(J).$$

We can conclude that (X, u, p) satisfies (C1–2), making it a $\mathfrak{U}$-counting system. ■

Remark 101. Theorem 97 and Theorem 100 prove that either of our universal properties is sufficient to define the stages of $\mathfrak{V}$, provided we restrict the set of destination to transitive elements of $\mathfrak{U}$ in the case of the $\mathfrak{U}$-counting system.

Lemma 102. *Let α be an ordinal. Then $\mathcal{V}_\alpha \in \mathfrak{V}$, and $\mathsf{rank}\,\mathcal{V}_\alpha = \alpha$.*

Proof. $\mathcal{V}_\alpha \in \mathsf{P}\mathcal{V}_\alpha = \mathcal{V}_{\alpha^+} \subseteq \mathfrak{V}$ and so $\mathcal{V}_\alpha \in \mathfrak{V}$. Clearly $\mathsf{rank}\,\mathcal{V}_\alpha \leqslant \alpha$. If $\mathsf{rank}\,\mathcal{V}_\alpha \neq \alpha$ then by Lemma 94, $\mathcal{V}_{\mathsf{rank}\,\mathcal{V}_\alpha} \subset \mathcal{V}_\alpha$, which is impossible since $\mathcal{V}_\alpha \subseteq \mathcal{V}_{\mathsf{rank}\,\mathcal{V}_\alpha}$. ■

Lemma 103. *$\mathcal{V}_\alpha$ is transitive for each ordinal α.*

Proof. Consider any $X \in \mathcal{V}_\alpha$. By Lemma 93, $X \in \mathcal{V}_{\alpha^+} = \mathsf{P}\mathcal{V}_\alpha$, and thus $X \subseteq \mathcal{V}_\alpha$. ■

Theorem 104. *$\mathfrak{V}$ is a universe satisfying (U4).*

Proof. We need to prove that (U1,3,4) hold for $\mathfrak{V}$ (recall that (U4) $\Rightarrow$ (U0,2)).

U1. Consider any $X \in \mathfrak{V}$. Then $X \in \mathcal{V}_\alpha$ for some $\alpha \in \mathcal{O}$. By Lemma 103, $X \subseteq \mathcal{V}_\alpha \subseteq \mathfrak{V}$.

U3. Consider any $X \in \mathfrak{V}$ and any function $f\colon X \to \mathfrak{V}$. Then $X \in \mathcal{V}_\alpha$ for some $\alpha \in \mathcal{O}$. Consider the composite $g = \mathsf{rank} \circ f$. Since $X \in \mathcal{V}_\alpha \in \mathfrak{U}$, by (U1), $X \in \mathfrak{U}$. Then $gX \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$ (Lemma 34). Then, by (O2), the incremented join $\gamma = \forall\!\!\!\vee\, gX$ exists. By Lemma 94, it is clear that $fX \subset \mathcal{V}_\gamma$. By Lemma 93, $\mathcal{V}_\gamma$ is a transitive set and so $\bigcup fX \subseteq \mathcal{V}_\gamma$. Then $\bigcup fX \in \mathsf{P}\mathcal{V}_\gamma = \mathcal{V}_{\gamma^+}$, and thus $\bigcup fX \in \mathfrak{V}$.

U4. Consider $X \in \mathfrak{V}$. Then $X \in \mathcal{V}_\alpha$ for some $\alpha \in \mathcal{O}$, and thus $X \subseteq \mathcal{V}_\alpha$ by transitivity of $\mathcal{V}_\alpha$. Then $\mathsf{P}X \subseteq \mathsf{P}\mathcal{V}_\alpha$, and thus $\mathsf{P}X \in \mathsf{PP}\mathcal{V}_\alpha = \mathcal{V}_{\alpha^{++}} \subseteq \mathfrak{V}$.

We can conclude that $\mathfrak{V}$ is a universe satisfying (U4). ■

Note that when $\mathfrak{V}$ is empty or has an infinite element, the following lemma already holds by Theorem 104 and Lemma 99.

Lemma 105. *If $X \in \mathfrak{V}$, then $X^T \in \mathfrak{V}$.*

Proof. If $X \in \mathfrak{V}$, then $X \in \mathcal{V}_\alpha$ for some ordinal α. By Lemma 103, $\mathcal{V}_\alpha$ is a transitive set. Therefore, the transitive closure X^T of X is a subset of $\mathcal{V}_\alpha$ and hence an element of $\mathcal{V}_{\alpha^+}$. By Lemma 102, $\mathcal{V}_{\alpha^+} \in \mathfrak{V}$. By transitivity of $\mathfrak{V}$ (Theorem 104), $X^T \in \mathfrak{V}$. ■

Lemma 106. $\operatorname{rank} X < \alpha$ *for any ordinal* α *and* $X \in \mathcal{V}_\alpha$.

Proof. Suppose $X \in \mathcal{V}_\alpha$. If α is a successor ordinal $\alpha = \beta^+$, then $X \subseteq \mathcal{V}_\beta$, and so $\operatorname{rank} X \leqslant \beta < \alpha$. If α is a limit ordinal, then $X \in \mathcal{V}_\beta$ for some $\beta < \alpha$ and so once again $\operatorname{rank} X \leqslant \beta < \alpha$ (note that by Lemma 103, $X \in \mathcal{V}_\beta$ implies $X \subseteq \mathcal{V}_\beta$). ■

Lemma 107. *Every nonempty set* $X \in \mathfrak{V}$ *has an element* x *such that* $X \cap x = \varnothing$.

Proof. We can take x to be an element of X such that $\operatorname{rank} x$ is minimal. Since, by Lemma 106, all elements of x have a rank strictly smaller than the rank of x, we will have $X \cap x = \varnothing$. ■

A set X is said to be *well-founded* if, for every nonempty subset Y of the transitive closure X^T of X, there exists $x \in Y$ such that $x \cap Y = \varnothing$. Such x is called an $\in$-minimal element of Y. From Lemma 107 and Theorem 104 we get, at once, the following result.

Lemma 108. *Each element* X *of* $\mathfrak{V}$ *is well-founded.*

The 'axiom of foundation', due to von Neumann [14], states that every set is well-founded. In [18], Zermelo included this axiom as part of the Zermelo-Fraenkel axiomatic set theory. However, as remarked by many authors (see e.g. [15]), this axiom is not needed for most of the theory of ordinal numbers, and accordingly, we have not assumed it in this thesis.

By $\mathfrak{V}^*$ we denote the set of partial functions $f\colon \mathfrak{V} \to \mathfrak{V}$. Then the domain $\operatorname{dom} f$ of each $f \in \mathfrak{V}^*$ is a subset of $\mathfrak{V}$. For a fixed element $x \in \mathfrak{V}$, consider the function $A_x\colon \mathrm{P}_{\mathfrak{U}}\mathfrak{V}^* \to \mathfrak{V}^*$ which maps each $F \in \mathrm{P}_{\mathfrak{U}}\mathfrak{V}^*$ to the partial function $A_x(F)$ defined by:

$$\operatorname{dom} A_x(F) = \bigcup\{\operatorname{dom} f \mid f \in F\};$$
$$y \mapsto x \cup \{f(t) \mid [f \in F] \wedge [t \in y \cap \operatorname{dom} f]\}.$$

Consider, also, the function $T_x\colon \mathfrak{V}^* \to \mathfrak{V}^*$ which maps each $f \in \mathfrak{V}^*$ to the partial function $T_x(f)$ defined by:

$$\operatorname{dom} T_x(f) = \mathrm{P}\operatorname{dom} f;$$
$$y \mapsto x \cup \{f(t) \mid t \in y\}.$$

By applying transfinite recursion (Theorem 57) to the triple $(\mathfrak{V}^*, A_x, T_x)$, we get, for each $x \in \mathfrak{V}$, a function

$$\mathcal{O} \to \mathfrak{V}^*, \quad \alpha \mapsto (x+_\alpha).$$

Each $(x+_\alpha)$ is thus a partial function $\mathfrak{V} \to \mathfrak{V}$. We write the values of this function as $(x+_\alpha)(y) = x +_\alpha y$. In this notation, the recursive functions that define $x +_\alpha y$ (with x fixed, as before) are:

$$\begin{aligned} \mathsf{dom}(x+_{\alpha^+}) &= \mathsf{P}\,\mathsf{dom}(x+_\alpha), \\ x +_{\alpha^+} y &= x \cup \{x +_\alpha t \mid t \in y\}; && \text{(for each ordinal } \alpha) \\ \mathsf{dom}(x+_\gamma) &= \bigcup\{\mathsf{dom}(x+_\beta) \mid \beta < \gamma\}, \\ x +_\gamma y &= x \cup \{x +_\beta t \mid [\beta < \gamma] \wedge [t \in y \cap \mathsf{dom}(x+_\beta)]\}. && \text{(for each limit ordinal } \gamma) \end{aligned}$$

From this we see that for each ordinal α,

$$\mathsf{dom}(x+_\alpha) = \mathcal{V}_\alpha.$$

Theorem 109. *For any $x \in \mathfrak{V}$ and ordinals $\alpha < \beta$, we have $x +_\alpha y = x +_\beta y$ for all $y \in \mathcal{V}_\alpha$. As a consequence, $(x+_\alpha)$ is a proper subfunction of $(x+_\beta)$ if and only if $\alpha < \beta$.*

Proof. We prove this by strong transfinite induction on β. Suppose that, for some ordinal β and all ordinals $\alpha' < \beta' < \beta$, it holds that $x +_{\alpha'} y = x +_{\beta'} y$ for all $y \in \mathcal{V}_{\alpha'}$.

Now consider any ordinal $\alpha < \beta$, and any $y \in \mathcal{V}_\alpha$. Since α and β may each be either a successor or a limit ordinal, we consider the following four cases.

Case 1: $\alpha = \delta^+$ and $\beta = \sigma^+$ for some ordinals δ and σ. Then $\delta < \sigma < \beta$, and thus

$$\begin{aligned} x +_\alpha y &= x +_{\delta^+} y \\ &= x \cup \{x +_\delta t \mid t \in y\} \\ &= x \cup \{x +_\sigma t \mid t \in y\} && \text{(by induction assumption)} \\ &= x +_{\sigma^+} y \\ &= x +_\beta y. \end{aligned}$$

Case 2: $\alpha = \delta^+$ for some ordinal δ, and β is a limit ordinal. Then, since $t \in \mathcal{V}_\delta$ for all $t \in y$, and $\delta < \beta$,

$$\begin{aligned} x +_\alpha y &= x \cup \{x +_\delta t \mid t \in y\} \\ &= x \cup \{x +_\delta t \mid t \in y \cap \mathcal{V}_\delta\} \\ &\subseteq x \cup \{x +_\sigma t \mid [\sigma < \beta] \wedge [t \in y \cap \mathcal{V}_\sigma]\} \\ &= x +_\beta y. \end{aligned}$$

Furthermore, since $t \in \mathcal{V}_\delta \subset \mathcal{V}_\alpha = \mathsf{dom}(x+_\alpha)$ for all $t \in y$,

$$\begin{aligned} x +_\beta y &= x \cup \{x +_\sigma t \mid [\sigma < \beta] \wedge [t \in y \cap \mathcal{V}_\sigma]\} \\ &\subseteq x \cup \{x +_\alpha t \mid t \in y\} && \text{(by induction assumption)} \\ &= x +_\alpha y. \end{aligned}$$

So $x +_\alpha y = x +_\beta y$.

Case 3: α is a limit ordinal and $\beta = \sigma^+$ for some ordinal σ. Then $\alpha \leqslant \sigma < \beta$, and thus

$$\begin{aligned} x +_\alpha y &= x \cup \{x +_\delta t \mid [\delta < \alpha] \wedge [t \in y \cap \mathcal{V}_\delta]\} \\ &\subseteq x \cup \{x +_\sigma t \mid t \in y\} && \text{(by induction assumption)} \\ &= x +_{\sigma^+} y \\ &= x +_\beta y. \end{aligned}$$

Notice that each $t \in y$ is in the domain of $(x+_\delta)$ for some $\delta < \alpha$, and when it is, we have $x +_\delta t = x +_\sigma t$ by induction assumption. Then

$$\begin{aligned} x +_\beta y &= x +_{\sigma^+} y \\ &= x \cup \{x +_\sigma t \mid t \in y\} \\ &\subseteq x \cup \{x +_\delta t \mid [\delta < \alpha] \wedge [t \in y \cap \mathcal{V}_\delta]\} \\ &= x +_\alpha y. \end{aligned}$$

So $x +_\alpha y = x +_\beta y$.

Case 4: α and β are both limit ordinals. Then $\alpha < \alpha^+ < \beta$. By induction assumption and Case 2,

$$x +_\alpha y = x +_{\alpha^+} y = x +_\beta y.$$

Since all four cases give us $x +_\alpha y = x +_\beta y$, we can conclude by induction that this equality holds for all $x \in \mathfrak{V}$, all ordinals α and β such that $\alpha < \beta$, and all $y \in \mathcal{V}_\alpha$.

Thus, $(x+_\alpha)$ is a subfunction of $(x+_\beta)$ for all $\alpha < \beta$, and furthermore, since

$$\mathsf{dom}(x+_\alpha) = \mathcal{V}_\alpha \subset \mathcal{V}_\beta = \mathsf{dom}(x+_\beta),$$

$(x+_\alpha)$ is a proper subfunction of $(x+_\beta)$. From this we also get that, whenever $(x+_\alpha)$ is a proper subfunction of $(x+_\beta)$, we have $(x+_\beta) \nsubseteq (x+_\alpha)$ and thus $\beta \nleqslant \alpha$, i.e. $\alpha < \beta$. ■

Lemma 110. *For any $S \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$, the following holds:*

$$A_x(\{(x+_\alpha) \mid \alpha \in S\}) = \bigcup\{(x+_\alpha) \mid \alpha \in S\} = (x+_{\bigvee S}).$$

Proof. Let $S \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$. Then, by Lemma 95,

$$\begin{aligned}\operatorname{dom} A_x(\{(x+_\alpha) \mid \alpha \in S\}) &= \bigcup\{\operatorname{dom}(x+_\alpha) \mid \alpha \in S\} \\ &= \bigcup\{\mathcal{V}_\alpha \mid \alpha \in S\} \\ &= \mathcal{V}_{\bigvee S} \\ &= \operatorname{dom}(x+_{\bigvee S}).\end{aligned}$$

By Theorem 109,

$$[(x+_\beta) \subset (x+_\alpha)] \quad \Leftrightarrow \quad [\beta < \alpha] \quad \Leftrightarrow \quad [\operatorname{dom}(x+_\beta) \subset \operatorname{dom}(x+_\alpha)].$$

This means that the identities we established for the domains of these functions uniquely determine the values of these functions as well, giving us, as required,

$$A_x(\{(x+_\alpha) \mid \alpha \in S\}) = \bigcup\{(x+_\alpha) \mid \alpha \in S\} = (x+_{\bigvee S}). \qquad \blacksquare$$

Theorem 111. *Let $X = \{(x+_\alpha) \mid \alpha \in \mathcal{O}\}$. Then the function $f\colon \mathcal{O} \to X$ defined by $\alpha \mapsto (x+_\alpha)$ is an order isomorphism between $(\mathcal{O}, \leqslant)$ and $(X, \subseteq)$. This makes $(X, \subseteq)$ an ordinal system with T_x as its successor function, and $\bigvee F$ given simultaneously by $A_x(F)$ and $\bigcup F$ for each $F \in \mathsf{P}_{\mathfrak{U}}X$.*

Proof. To see that $(X, \subseteq)$ is an ordinal system, we prove that it is order-isomorphic (see Definition 21) to $(\mathcal{O}, \leqslant)$. Clearly, the function $f\colon \mathcal{O} \to X$ defined by $\alpha \mapsto (x+_\alpha)$ is surjective. Then by Theorem 109, f is an order isomorphism. We can conclude that $(X, \subseteq)$ is an ordinal system.

Since $(x+_{\alpha^+}) = T_x(x+_\alpha)$ and $(\mathcal{O}, \leqslant)$ and $(X, \subseteq)$ are order-isomorphic, T_x is the successor function of X. By Lemma 110, for each $S \in \mathsf{P}_{\mathfrak{U}}\mathcal{O}$,

$$\begin{aligned}\bigvee fS &= f(\bigvee S) \\ &= (x+_{\bigvee S}) \\ &= A_x(fS) \\ &= \bigcup fS.\end{aligned}$$

Since f is an order isomorphism, this is equivalent to the statement that $\bigvee F$ is given simultaneously by $A_x(F)$ and $\bigcup F$ for each $F \in \mathsf{P}_{\mathfrak{U}}X$. $\blacksquare$

We can now define a binary operation of *Tarski addition* on $\mathfrak{V}$ as follows: for any two $x, y \in \mathfrak{V}$,

$$\begin{aligned}x + y &= x +_{(\operatorname{rank} y)^+} y \\ &= x \cup \{x + t \mid t \in y\}.\end{aligned}$$

This is the form in which addition of sets was defined by Alfred Tarski at a 1955 meeting of the American Mathematical Society [9], but never published. This operation extends the usual addition of von Neumann ordinals to arbitrary sets in the cumulative hierarchy. This, and other facts about this addition, have been explored in [10]. In our context one can reobtain these facts almost verbatim. Among them is strict associativity, which we prove as follows.

Theorem 112. *Tarski addition is associative, i.e. for any $x, y, z \in \mathfrak{V}$, we have*

$$x + (y + z) = (x + y) + z.$$

Proof. We prove this by strong transfinite induction on α such that $z \in \mathcal{V}_\alpha$. Let $z \in \mathcal{V}_\alpha$, and suppose that, for all $t \in \mathcal{V}_\beta$ such that $\beta < \alpha$, it holds that $x + (y + t) = (x + y) + t$. Then

$$\begin{aligned}
x + (y + z) &= x + (y \cup \{y + t \mid t \in z\}) \\
&= x \cup \{x + s \mid s \in y \cup \{y + t \mid t \in z\}\} \\
&= x \cup \{x + s \mid s \in y\} \cup \{x + s \mid s \in \{y + t \mid t \in z\}\} \\
&= (x + y) \cup \{x + (y + t) \mid t \in z\} \\
&= (x + y) \cup \{(x + y) + t \mid t \in z\} && \text{(by induction assumption)} \\
&= (x + y) + z.
\end{aligned}$$

■

Bibliography

[1] Hilbert, D.: Über das Unendliche. *Mathematische Annalen*, vol. 95, no. 1, pp. 161–190, 1926.

[2] Dedekind, J.W.R.: *Was Sind und was Sollen die Zahlen?* Vieweg, Braunschweig, 1888.

[3] Peano, G.: *Arithmetices principia: Nova methodo exposita*. Fratres Bocca, 1889.

[4] Zermelo, E.F.F.: Untersuchungen über die Grundlagen der Mengenlehre (I). *Mathematische Annalen*, vol. 65, no. 2, pp. 261–281, 1908.

[5] Mac Lane, S.: *Categories for the Working Mathematician*, vol. 5 of *Graduate Texts in Mathematics*. 2nd edn. Springer, 1998.

[6] Bourbaki, N., Artin, M., Grothendieck, A. & Verdier, J.-L.: Théorie des topos et cohomologie étale des schémas: Séminaire de géométrie algébrique du Bois Marie (Tome 1), 1963–1964.

[7] Joyal, A. & Moerdijk, I.: *Algebraic Set Theory*, vol. 220. Cambridge University Press, 1995.

[8] Janelidze, Z. & van der Berg, I.: A Dedekind-style axiomatization and the corresponding universal property of an ordinal number system. *arXiv preprint:2006.12688*, 2020.

[9] Tarski, A.: The notion of rank in axiomatic set theory and some of its applications. In: *The June Meeting in Vancouver*, vol. 61 of *Bulletin of the American Mathematical Society*, p. 443. 1955.

[10] Kirby, L.: Addition and multiplication of sets. *Mathematical Logic Quarterly*, vol. 53, no. 1, pp. 52–65, 2007.

[11] Enderton, H.B.: *A Mathematical Introduction to Logic*. Academic Press, 1972.

[12] Kunen, H.K.: *Set Theory: an Introduction to Independence Proofs*. Elsevier Science Publishers, 1980.

[13] Fraenkel, A.H.: Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre. *Mathematische Annalen*, vol. 86, no. 3-4, pp. 230–237, 1922.

[14] Von Neumann, J.: Eine Axiomatisierung der Mengenlehre. *Journal für die Reine und Angewandte Mathematik*, vol. 1925, no. 154, pp. 219–240, 1925.

[15] Jech, T.J.: *Set Theory*. Springer, 2003.

[16] Davey, B.A. & Priestley, H.A.: *Introduction to Lattices and Order*. Cambridge University Press, 2002.

[17] Van Heijenoort, J.L.M.: *From Frege to Gödel: a Source Book in Mathematical Logic, 1879–1931*, vol. 9. Harvard University Press, 1967.

[18] Zermelo, E.F.F.: Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Mengenlehre. *Fundamenta Mathematicæ*, 1930.

[19] Devlin, K.J.: *The Joy of Sets: Fundamentals of Contemporary Set Theory*. Springer, 2012.

[20] Ebbinghaus, H.-D.: *Ernst Zermelo: An Approach to His Life and Work*. Springer Science & Business Media, 2007.

[21] Lawvere, F.W.: An elementary theory of the category of sets. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 52, no. 6, pp. 1506–1511, 1964.

[22] Fraenkel, A.H.: Abstract set theory. *Bulletin de la Société Mathématique de France*, vol. 59, pp. 584–585, 1953.

[23] Williams, N.H.: On Grothendieck universes. *Composito Mathematica*, vol. 21, no. 1, pp. 1–3, 1969.

[24] Ore, Ø.: Galois connexions. *Transactions of the American Mathematical Society*, vol. 55, no. 3, pp. 493–513, 1944.

[25] Von Neumann, J.: Zur Einführung der transfiniten Zahlen. *Acta Litterarum ac Scientiarum Regiae Universitatis Hungaricae Francisco-Josephinae, Sectio Scientiarum Mathematicarum*, vol. 1, pp. 199–208, 1923.

[26] Alexandroff, P.S.: Diskrete Räume. *Математический Сборник (Matematicheskii Sbornik)*, vol. 2, no. 3, pp. 501–519, 1937.

[27] Johnstone, P.T.: *Stone spaces*, vol. 3. Cambridge university press, 1982.

[28] McCord, M.C. *et al.*: Singular homology groups and homotopy groups of finite topological spaces. *Duke Mathematical Journal*, vol. 33, no. 3, pp. 465–474, 1966.

[29] Hrbáček, K. & Jech, T.J.: *Introduction to Set Theory, Revised and Expanded*. Chemical Rubber Company Press, 1999.